



Università degli Studi di Napoli

"Federico II"

Facoltà di Scienze MM.FF.NN

DOTTORATO DI RICERCA

IN

BIOLOGIA AVANZATA

Indirizzo: Sistematica molecolare

XXII ciclo

**"I geni MADS-box e l'evoluzione molecolare
dello sviluppo florale nelle Orchidaceae"**

Tutore

Prof.ssa Serena Aceto

Coordinatore

Prof. Luciano Gaudio

Dottoranda

Dott.ssa Carmela Cantone

Anno accademico 2008-2009

COMUNICAZIONI A CONGRESSO

1) Cantone C., Sica M., Gaudio L., Aceto S. (2007). Promoter isolation, intron variability, and micro-RNA putative target analysis at the *OrcPI* locus in orchids. *FISV IX National Congress Acta*, September 26-29, 2007 Riva del Garda, Italy. PMS.7 P13.

2) Cantone C., Sica M., Gaudio L., Aceto S. (2008). Genomic organization and expression pattern of the *OrcPI* locus in *Orchis italica* (Orchidaceae). *FISV X National Congress Acta*, September 24-27, 2008 Riva del Garda, Italy. T.8 P15 .

3) Cantone C., Aceto S., Chiaiese P., Ruotolo G., Gaudio L. (2009). Characterization of the promoter region of *OrcPI*, a MADS-box gene involved in flower development in orchids. SIBBM 2009 Frontiers in Molecular Biology, June 4-6, 2009 Naples, Italy, P2.1, pag 84.

4) Cantone C., Gaudio L., Aceto S. (2009). *GLO*-like and *DEF*-like MADS-box genes in orchids: copy number and evolutionary implications. *FISV XI National Congress Acta*, September 23-25, 2009 Riva del Garda, Italy. D09.03.

PUBBLICAZIONI IN EXTENSO

1) Cantone C., Sica M., Gaudio L., Aceto S. (2009). The *OrcPI* locus: genomic organization, expression pattern, and noncoding regions variability in *Orchis italica* (Orchidaceae) and related species. *Gene*, 434: 9-15. DOI: 10.1016/j.gene.2008.12.015.

2) Aceto S., **Cantone C.**, Chiaiese P., Ruotolo G., Sica M., Gaudio L. (2009). Isolation and phylogenetic footprinting analysis of the promoter region of the floral homeotic gene *OrcPI* from *Orchis italica* (Orchidaceae). *Journal of Heredity*. DOI: 10.1093/jhered/esp082.

Indice

<u>Introduzione</u>	pag.	5
I geni MADS-box e lo sviluppo fiorale		5
La storia evolutiva dei geni MADS-box		10
<i>Diversificazione funzionale dei geni MADS-box e lo sviluppo fiorale</i>		12
Diversità fiorale nelle angiosperme basali		13
Diversità fiorale delle orchidee		15
<i>Organi riproduttivi: ruolo dei geni omeotici fiorali</i>		18
<i>L' "Orchid code"</i>		21
<i>Evoluzione dei geni DEF-like nelle orchidee</i>		22
<u>Scopo del lavoro</u>		26
<u>Materiali</u>		28
<u>Metodi</u>		28
Allestimento e <i>screening</i> di una mini-libreria di cDNA di geni MADS-box in <i>Orchis italica</i>		28
Identificazione delle regioni 5'UTR dei geni <i>OrcPI</i> e <i>OrcPI2</i> in <i>O. italica</i>		33
Identificazione del promotore del gene <i>OrcPI</i> in <i>O. italica</i>		33
Analisi funzionale del promotore del gene <i>OrcPI</i>		40
Caratterizzazione strutturale del locus <i>OrcPI</i> in <i>O. italica</i> e in altre specie di <i>Orchidoideae</i>		44
<i>Promotore</i>		44
<i>Esoni e introni</i>		47
Analisi dell'espressione del gene <i>OrcPI</i> in <i>O. italica</i>		50

Caratterizzazione strutturale del locus <i>OrcPI2</i>	
in <i>O. italica</i> e in altre specie di <i>Orchidoideae</i>	51
Amplificazione, clonaggio e sequenziamento di regioni	
genomiche dei geni <i>GLO-like</i> di <i>Habenaria radiata</i>	54
Analisi delle sequenze	57
<u>Risultati e Discussione</u>	63
Identificazione dei geni MADS-box in <i>Orchis italica</i>	63
Analisi dei geni <i>AG-like</i> in <i>O. italica</i>	63
Analisi dei geni <i>DEF-like</i> in <i>O. italica</i>	66
Analisi dei geni <i>GLO-like</i> in <i>O. italica</i>	72
Il locus <i>OrcPI</i>	79
Analisi del promotore del gene <i>OrcPI</i>	81
Caratterizzazione strutturale e variabilità	
nucleotidica del gene <i>OrcPI</i>	84
Analisi dell'espressione del gene <i>OrcPI</i> in <i>O. italica</i>	89
<u>Bibliografia</u>	94

Introduzione

I geni MADS-box e lo sviluppo florale

L'enorme diversificazione delle angiosperme è considerata un classico esempio di radiazione adattativa, contrassegnata dalla comparsa di strutture riproduttive uniche: i fiori (Friis et al., 2006). Queste strutture riproduttive altamente specializzate mostrano un'enorme diversità morfologica strettamente correlata al successo evolutivo delle angiosperme (Lawton-Rauh et al., 2000).

Le modificazioni dei meccanismi che controllano lo sviluppo sono considerate responsabili della maggior parte della diversità morfologica che si osserva in natura e l'approccio genetico-molecolare è di fondamentale importanza per comprendere l'origine e la diversificazione delle strutture fiorali (Purugganan et al., 1998; Baum et al., 1999; Carroll, 2000).

La maggior parte delle differenze strutturali e funzionali manifestate dagli organi fiorali sono riconducibili all'attività di geni regolativi (geni selettori omeotici) coinvolti nel processo di fioritura, che consta di cinque stadi (Okada e Shimura, 1994):

- formazione del meristema dell'infiorescenza;
- determinazione dell'identità del meristema florale;
- determinazione del numero e della disposizione degli organi fiorali;
- determinazione dell'identità degli organi;
- maturazione degli organi fiorali.

La funzione che i geni omeotici svolgono durante la morfogenesi florale è stata descritta, per la prima volta, nel modello genetico "ABC"; esso, sulla base dell'analisi di mutanti delle specie modello *Arabidopsis thaliana* e *Antirrhinum majus*, ipotizza l'esistenza di tre funzioni regolative: la funzione A controlla lo sviluppo dei sepali, le funzioni A e B insieme dirigono la formazione dei petali, le funzioni B e C specificano gli stami e la funzione C da sola determina il fenotipo carpelli; le funzioni A e C, inoltre, si inibiscono a vicenda (Coen e Meyerowitz, 1991; Theissen, 2001).

Successivamente, il modello "ABC" è stato rivisitato, con l'aggiunta di altre due classi funzionali, note come D e E: la prima determina lo sviluppo degli ovuli e la seconda è richiesta per lo sviluppo di tutti gli organi fiorali (Figura 1) (Pelaz et al., 2001; Pinyopich et al., 2003).

I meccanismi molecolari delle interazioni che si instaurano tra queste diverse classi funzionali sono descritte nel modello del "quartetto florale", il quale suggerisce che l'identità degli organi fiorali è determinata da complessi tetramericici ("quartetti") costituiti dalla differente combinazione di quattro proteine omeotiche (Figura 2) (Theißen, 2001).

I loci omeotici coinvolti nel modello "ABCDE" appartengono alla famiglia genica MADS-box, fatta eccezione per il gene *APETALA2* (*AP2*) di funzione A in *A. thaliana* e i suoi ortologhi.

I geni MADS-box codificano fattori di trascrizione di importanza fondamentale nei processi di sviluppo e di trasduzione del segnale nella maggior parte degli eucarioti (Messenguy e Dubois, 2003); essi condividono una sequenza di DNA molto conservata, lunga circa 180 bp, la regione MADS-box, la quale codifica il dominio di legame al DNA. La parola MADS è l'acronimo dei loci in cui è stato identificato per la prima volta il dominio MADS-box: **MINI CHROMOSOME MAINTENANCE 1** (*MCM1*) in *Saccharomyces cerevisiae*, **AGAMOUS** (*AG*) in *Arabidopsis thaliana*, **DEFICIENS** (*DEF*) in *Antirrhinum majus* e **SERUM RESPONSE FACTOR** (*SRF*) in *Homo sapiens* (Becker e Theißen, 2003).

I geni MADS-box possono essere suddivisi in due super-gruppi, noti come tipo I e tipo II, originati da una duplicazione genica avvenuta prima della divergenza di piante e animali (Figura 3) (Jongmin et al., 2004).

I fattori di trascrizione MADS-box, responsabili dell'ontogenesi florale appartengono al tipo II e presentano un'organizzazione proteica modulare, definita di tipo MIKC, comprendente quattro regioni principali: il dominio MADS (M), il dominio *Intervening* (I), il dominio cheratina simile (K) e il dominio C-terminale (C). Il dominio MADS, la cui sequenza è molto conservata, interagisce con il DNA ed è coinvolto nella dimerizzazione delle proteine; il dominio I, meno conservato, rende specifica la dimerizzazione; il dominio K media l'interazione tra le proteine e il dominio C, con

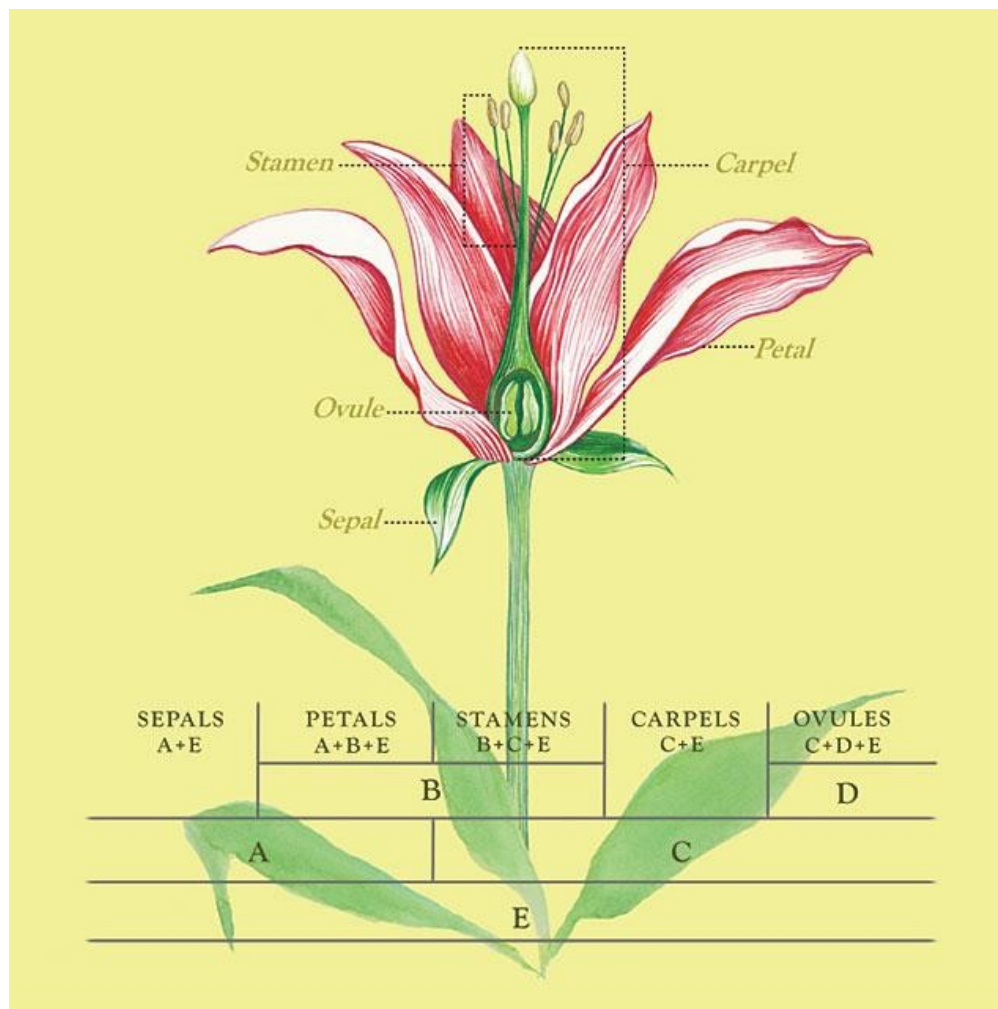


Figura 1 - Rappresentazione grafica del modello "ABCDE". (Da Soltis et al., Annals of Botany 2007)

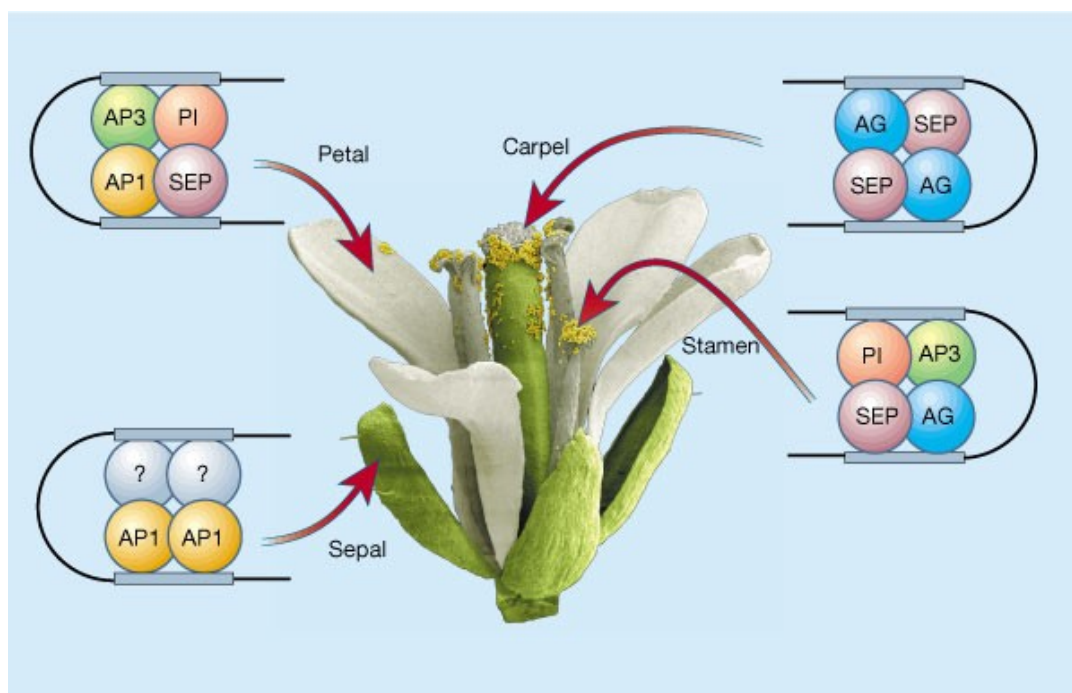


Figura 2 - Rappresentazione grafica del modello del "quartetto florale" (AP3, APETALA 3; PI, PISTILLATA; AP1, APETALA 1; SEP, SEPALLATA; AG, AGAMOUS). (Da Theißen e Saedler, Nature 2001)

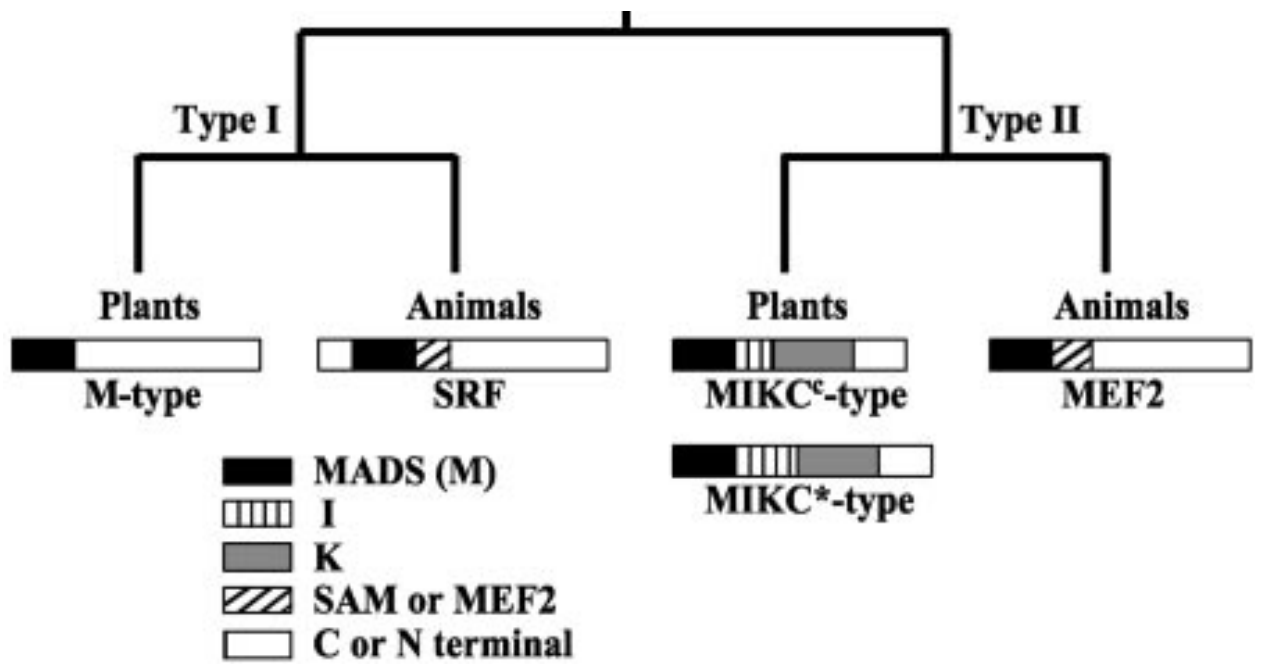


Figura 3 - Rappresentazione grafica della diversificazione strutturale dei domini funzionali dei geni MADS-box di tipo I e di tipo II, dopo duplicazione. (Da Nam et al., PNAS 2004)

sequenza molto variabile, è coinvolto nella formazione di complessi multimerici tra i fattori di trascrizione (Figura 4) (Kaufmann et al., 2005).

Nelle dicotiledoni, i geni MIKC di tipo II possono essere suddivisi in quattro famiglie i cui membri esibiscono profili di espressione e funzioni molto simili:

- *AP1/SQUA* (*APETALA1/SQUAMOSA*) (funzione A)
- *AP3/PI* (*APETALA3/PISTILLATA*) (funzione B)
- *AG/AGL11* (*AGAMOUS*) (funzione C)
- *SEP* (*SEPALLATA*) (funzione E)

L'evoluzione fiorale è governata, quindi, da un complesso e fine programma genetico i cui protagonisti sono i loci MADS-box di tipo MIKC; per tale motivo, la comprensione degli aspetti fondamentali che caratterizzano le innovazioni morfologiche nelle piante presuppone la conoscenza della storia evolutiva dei geni MADS-box (Rijpkema et al., 2007).

Storia evolutiva dei geni MADS-box

Le duplicazioni rappresentano un'importante fonte di diversificazione genica, la quale sembra instaurarsi attraverso due fenomeni: la sub-funzionalizzazione, processo nel quale le funzioni o i profili di espressione sono ripartiti tra i geni duplicati, e la neo-funzionalizzazione, con la quale il gene duplicato acquisisce una nuova funzione o nuovo profilo di espressione (Vandenbussche et al., 2003; Malcomber e Kellog, 2005). Gli eventi duplicativi e la successiva diversificazione funzionale dei geni duplicati, in particolare dei geni coinvolti nei meccanismi di sviluppo, sembra aver contribuito profondamente all'evoluzione di nuove caratteristiche morfologiche (Nei, 2005; Nei e Rooney, 2005; Freeling e Thomas, 2006).

Nelle piante la famiglia genica MADS-box si è ampliata ed evoluta attraverso duplicazioni successive seguite spesso da divergenza funzionale e strutturale dei geni (De Bodt et al., 2006), con corrispondenza temporale tra le duplicazioni dei geni MADS-box ed eventi importanti che hanno contrassegnato l'evoluzione delle angiosperme (Soltis et al., 2007).

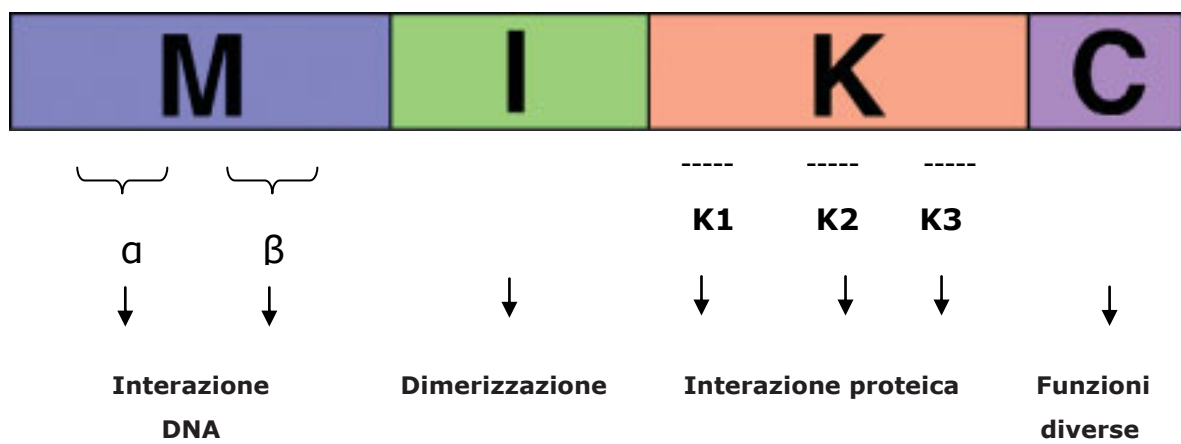


Figura 4 - Struttura modulare di tipo MIKC delle proteine MADS-box di tipo II.

Come detto precedentemente, i geni MADS-box coinvolti nello sviluppo del fiore possono essere raggruppati in diverse famiglie: *AP1/SQUA* (funzione A), *AP3/PI* (funzione B), *AG/AGL11* (funzione C), *SEP* (funzione E). Una serie di duplicazioni avvenute prima dell'origine delle angiosperme ha dato origine alle linee *AP3* e *PI* nella famiglia *AP3/PI*, alle linee *AG* e *AGL11* nella famiglia *AG/AGL11* e alle linee *AGL2/3/4* e *AGL9* nella famiglia *SEP*; un'altra serie di duplicazioni avvenute prima dell'origine delle dicotiledoni ha dato origine alle linee *euAP1*, *euFUL* e *AGL79* nella famiglia *AP1/SQUA*, *euAP3* and *TM6* nella linea *AP3*, *euAG* e *PLE* nella linea *AG* e *AGL2/4*, *AGL3* e *FBP9* nella linea *AGL2/3/4* (Kramer et al., 1998; Litt e Irish, 2003; Kim et al., 2004; Kramer et al., 2004; Zahn et al., 2005; Shan et al., 2007; Xu e Kong, 2007).

Recentemente sono stati descritti i possibili siti dei domini K e C-terminale dei geni MADS-box responsabili della diversificazione strutturale e funzionale delle quattro famiglie descritte (Shan et al., 2009). Come specificato in precedenza, i domini K e C-terminale sono coinvolti nella formazione dei complessi multi-proteici; è stato ipotizzato che lo scenario evolutivo descritto (duplicazione seguita da rilassamento delle forze selettive) abbia influenzato importanti caratteristiche fenotipiche delle piante, come lo sviluppo fiorale, attraverso l'instaurarsi di nuove interazioni proteiche e, di conseguenza, di nuove funzioni regolative (Hernandez-Hernandez et al., 2007; Rijpkema et al., 2007).

Diversificazione funzionale dei geni MADS-box di tipo B e sviluppo fiorale

La famiglia genica di classe B rappresenta un tipico esempio di come duplicazioni seguite da diversificazione funzionale possano influenzare profondamente l'origine e la diversificazione degli organismi.

I geni di funzione B sono suddivisi in due sotto-famiglie, note come *PISTILLATA/GLOBOSA-like* (*PI/GLO-like* o, più comunemente *GLO-like*) e *APETALA3/DEFICIENS-like* (*AP3/DEF-like* o, più comunemente *GLO-like*) dopo la loro caratterizzazione in *A. thaliana* e *A. majus*, rispettivamente. Come descritto nel modello "ABCDE", le

proteine di funzione B determinano lo sviluppo di petali e stami formando eterodimeri GLO/DEF e complessi tetramericici con l'eterodimero AP1/SEP (per la determinazione del fenotipo petali) o con l'eterodimero AG/SEP (per la specificazione degli stami) (Messenguy e Dubois, 2003). L'origine dell'eterodimerizzazione obbligata delle proteine B osservata nelle dicotiledoni sembra essersi evoluta da una situazione di omodimerizzazione (presente nelle gimnosperme) e uno stato transitorio di omo-eterodimerizzazione (Figura 5) (Davies et al., 1996; Riechmann et al., 1996; Winter et al., 2002). L'eterodimerizzazione sembra essersi stabilita in seguito alla diversificazione funzionale dei geni di classe B, guidata da selezione positiva in particolari siti del domini K e C, dopo le due principali duplicazioni: la prima (la duplicazione *AP3-PI*) avvenuta circa 260 milioni di anni fa, appena prima della divergenza tra angiosperme e gimnosperme, e la seconda (la duplicazione *euAP3-TM6*) datata 90 milioni di anni fa, coincidente con l'origine delle dicotiledoni (Figura 5) (Hernandez-Hernandez et al., 2007). Si ritiene che l'evoluzione dei geni di tipo B abbia avuto un ruolo centrale nell'origine e nella diversificazione delle angiosperme. In particolare, la diversificazione funzionale dei loci di classe B dopo la prima duplicazione ha contribuito all'origine di nuove strutture morfologiche (gli organi fiorali) e quindi all'origine dei fiori; la diversificazione funzionale dopo la seconda duplicazione è stata invece fondamentale per lo sviluppo florale caratteristico delle dicotiledoni (perianzio con organi differenziati) (Hernandez-Hernandez et al., 2007).

Diversità florale nelle angiosperme basali

Rispetto alle dicotiledoni, le angiosperme basali esibiscono una maggiore diversificazione della struttura e dell'organizzazione florale. Ad esempio, alcuni *taxa* hanno un perianzio non differenziato (*Amborellaceae*), in altri il perianzio è formato da petali e sepali morfologicamente distinti (*Annonaceae* e *Aristolochiaceae*); inoltre, alcune famiglie sono caratterizzate da una insolita proliferazione di organi fiorali ed altre dall'evoluzione di particolari strutture [ad esempio, la presenza

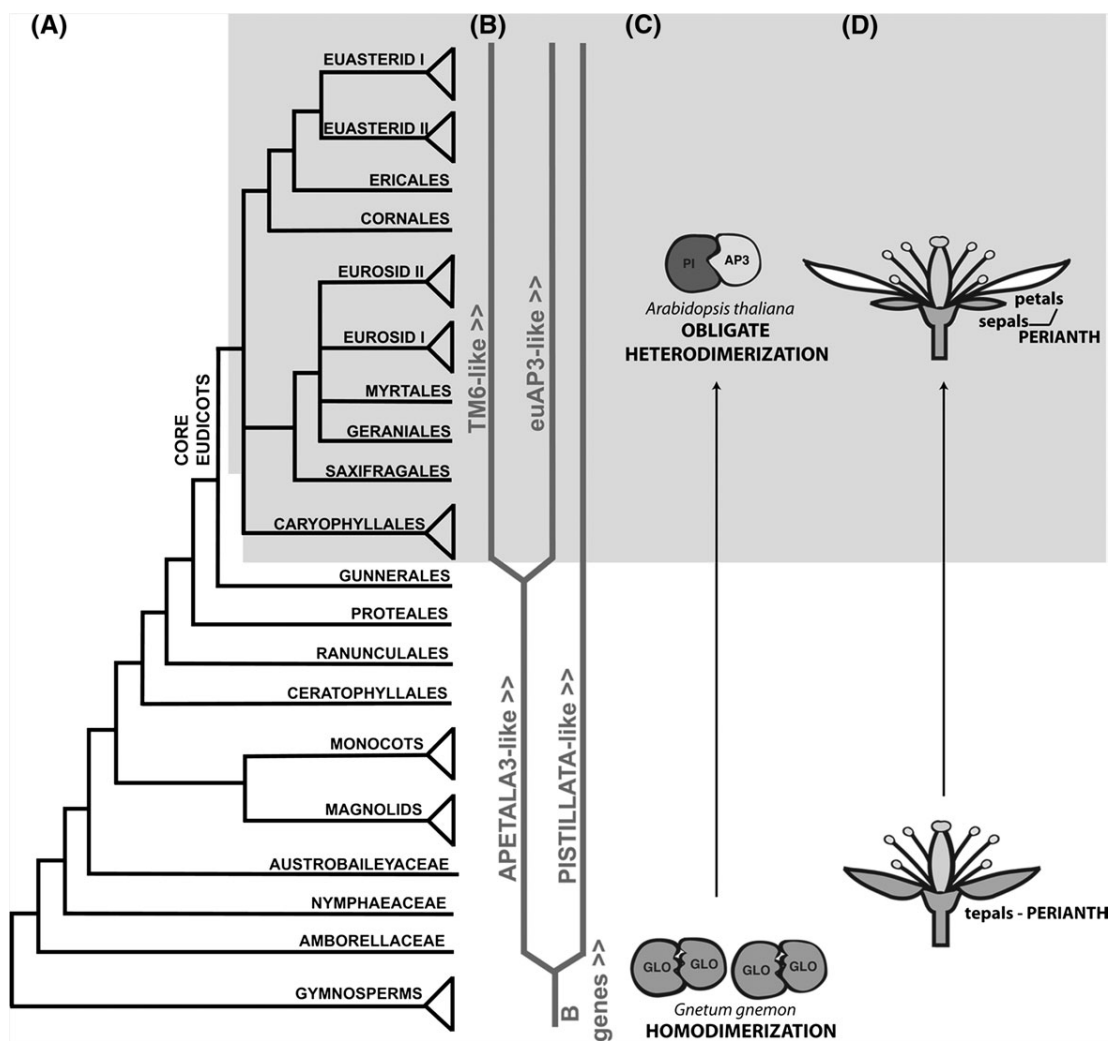


Figura 5 - **(A)** Rappresentazione grafica della filogenesi delle angiosperme. **(B)** Storia evolutiva dei geni MADS-box di classe B. **(C)** Evoluzione dell'interazione proteica tra le proteine AP3 e PI di *A. thaliana*. **(D)** Diversificazione florale delle angiosperme. (Da Hernandez-Hernandez et al., Mol. Biol. Evol. 2007)

di un centinaio di stami in *Tambourissa* (Manimiaceae) e l'evoluzione della caliptra in *Eupomatia* (Eupomatiaceae)] (Figura 6) (Kim et al., 2005; Chanderbali et al., 2006).

La funzione dei geni MADS-box è strettamente correlata al loro profilo di espressione ed è stato ipotizzato che la diversificazione florale osservata durante l'evoluzione delle angiosperme sia legata a cambiamenti del profilo di espressione dei geni MADS-box. L'ipotesi formulata è stata successivamente supportata da prove sperimentali (Soltis et al., 2007). Ad esempio, studi condotti su *Amborellaceae*, *Nymphaeaceae* ed altre angiosperme basali caratterizzate da un perianzio non differenziato hanno evidenziato che l'espressione dei geni di classe B e C non è confinata solo al secondo e al terzo verticillo (profilo tipico dei geni di classe B nelle dicotiledoni) e al terzo e quarto verticillo (profilo tipico dei geni di classe C nelle dicotiledoni), ma si estende anche agli organi esterni del perianzio (primo verticillo) (Buzgo et al., 2004; Kim et al., 2005; Chanderbali et al., 2006). Sulla base di queste osservazioni è stato sviluppato il modello "*fading borders*", secondo cui la formazione di organi floreali con morfologia simile è sottesa da un gradiente del livello di espressione dei geni MADS-box lungo il meristema florale: ai margini di un dominio di un gene la sua espressione si sovrappone con quella dei geni adiacenti, dando origine ad organi floreali con caratteristiche morfologiche non ben distinte (Figura 7) (Theißen et al., 2001; Soltis et al., 2007).

E' stato, inoltre, ipotizzato che il tipo di espressione descritto nel modello "*fading borders*" sia una condizione ancestrale; successivamente, con l'evoluzione di nuove interazioni proteina-proteina, di nuove funzioni regolative e in particolare del controllo autoregolativo, si è stabilito un profilo di espressione più ristretto, tipico delle dicotiledoni, il quale è ben descritto nel modello "ABCDE" (Theißen et al., 2001).

Diversità florale delle orchidee

Le *Orchidaceae*, con quasi 30.000 specie, rappresentano una delle più grandi famiglie delle angiosperme, suddivisa in cinque



Figura 6 – Fotografie di fiori di diverse angiosperme basali. **a)** *Alborella trichopoda*; **b)** *Nuphar advena*; **c)** *Illicium floridanum*; **d)** *Asinina longifolia*; **e)** *Magnolia grandiflora*; **f)** *Eupomatia bennetti*

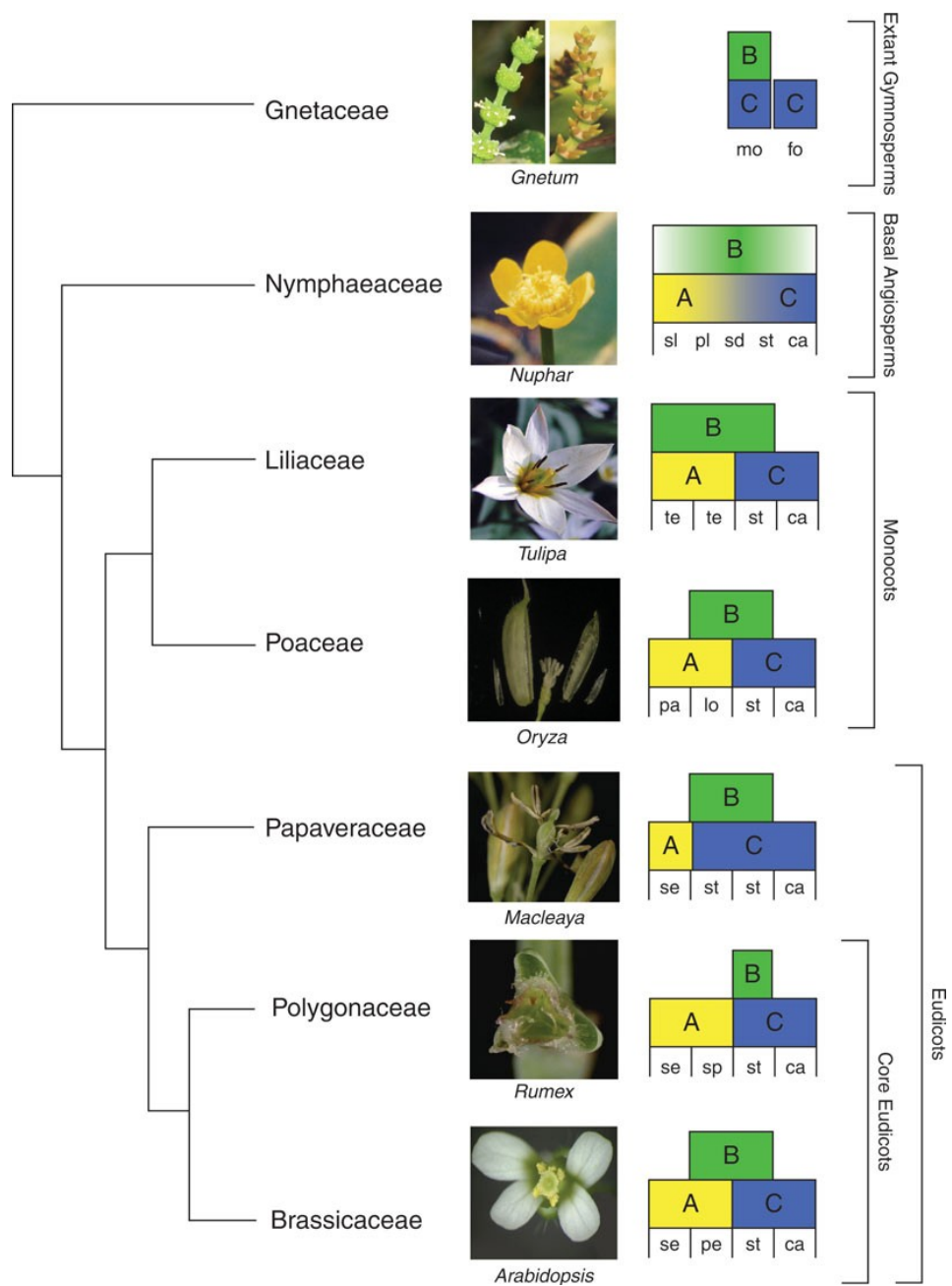


Figura 7 - Evoluzione del modello ABC. (Da Theissen e Melzer, Annals of Botany 2007)

sottofamiglie: *Apostasioideae*, *Cypripedioideae*, *Vanilloideae*, *Orchidoideae* ed *Epidendroideae* (Figura 8) (Ramirez et al., 2007).

I fiori delle orchidee, caratterizzati da una marcata diversificazione, sono considerati l'elemento chiave della loro radiazione adattativa (Mondragon-Palomino e Theißen, 2008). Il fiore delle orchidee ha una struttura zigomorfica con un perianzio costituito da tre tipi di organi: tre tepali esterni nel primo foglietto, due tepali interni laterali e uno interno mediano (labello) nel secondo foglietto (Figura 9) (Rudall e Bateman, 2002). I tepali esterni e quelli interni laterali hanno colore e forma simili, mentre il labello si distingue per dimensioni, forma e colore e la sua spettacolare diversità morfologica è considerata il risultato della co-evoluzione con gli impollinatori.

Un'altra peculiarità dei fiori delle orchidee è la fusione degli organi riproduttivi (stami/antere e pistilli/stigma) in un'unica struttura: la colonna (Rudall e Bateman, 2002). La variabilità della dimensione e della forma, la presenza o assenza di appendici, la conformazione, la posizione e il numero delle antere e le caratteristiche dei pollinii rendono la colonna un organo florale molto complesso ed altamente specializzato.

La particolarità del fiore delle orchidee da sempre ha suscitato grande interesse e sono stati condotti numerosi studi per comprendere la sua origine e diversificazione, tra cui molti focalizzati sui meccanismi ecologici (caratteristiche fiorali, impollinazione, comportamento degli impollinatori) (Johnson et al., 1998; Cozzolino e Widmer, 2005; Jersakova et al., 2006). Recentemente, grazie all'aumento delle conoscenze sulla base molecolare dello sviluppo del fiore, sono state condotte molte analisi genetico-molecolari per comprendere la morfogenesi florale delle *Orchidaceae* ed i risultati ottenuti hanno reso possibile sviluppare una teoria sull'evoluzione del fiore in questa famiglia.

Organi riproduttivi: ruolo dei geni omeotici fiorali

La biologia riproduttiva delle orchidee è un aspetto molto interessante di questa famiglia. Come descritto precedentemente, la fusione degli organi riproduttivi in un'unica struttura, la

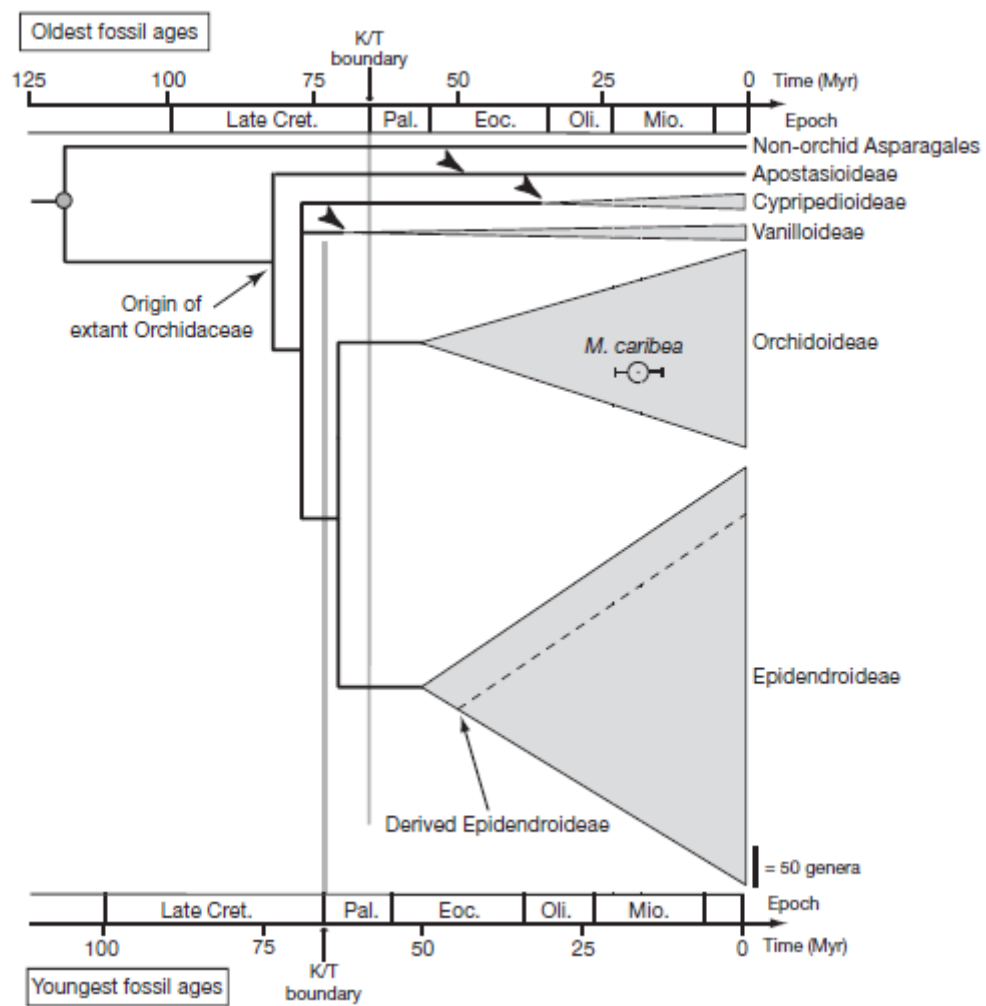
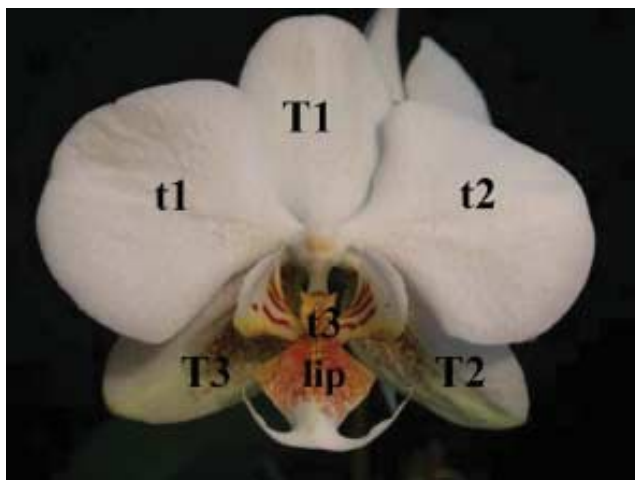


Figura 8 - Cladogramma delle *Orchidaceae*. (Da Ramirez et al., Nature 2007)

(A)



(B)

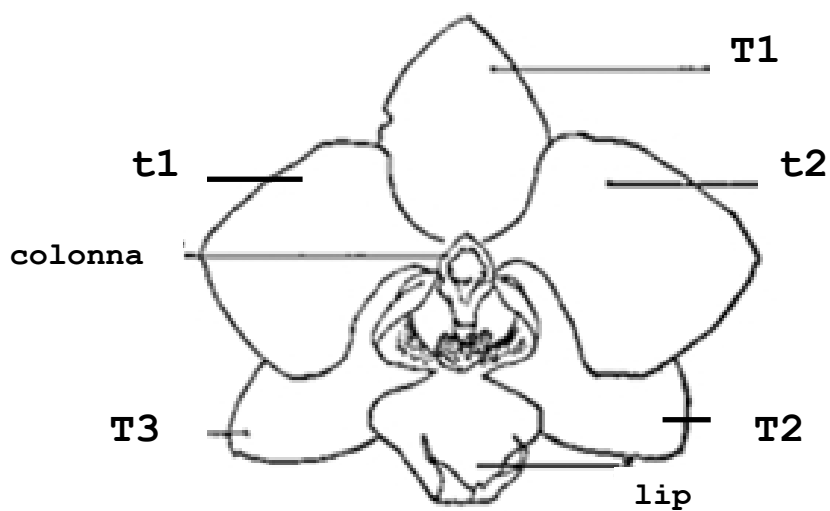


Figura 9 - Morfologia del fiore di *Phalaenopsis*. A) Fotografia del fiore di *Phalaenopsis*; B) diagramma degli organi fiorali di *Phalaenopsis*. T1, T2, T3 - tepali esterni; t1, t2 - tepali interni laterali; lip - labello.

colonna, rappresenta una delle peculiarità del fiore delle orchidee; un'altra caratteristica riguarda lo sviluppo degli ovuli. Nella maggior parte dei fiori delle angiosperme, l'ovario e gli ovuli maturano durante l'ontogenesi fiorale, mentre nelle orchidee lo sviluppo dell'ovario e degli ovuli è strettamente correlato all'impollinazione: due o tre mesi dopo l'impollinazione, gli ovuli iniziano a svilupparsi e solo quando sono completamente maturi avviene la fecondazione (Zhang et al., 1993; Nadeau et al., 1996).

Dato che i geni di classe B, C e D specificano l'identità delle strutture riproduttive dei fiori delle angiosperme, per analizzare la formazione della colonna e chiarire l'inusuale sviluppo degli ovuli sono stati isolati e analizzati i profili di espressione di tali geni in diverse orchidee (Tsai et al., 2005; Skipper et al., 2006; Song et al., 2006; Xu et al., 2006). In *Phalaenopsis* e in *Dendrobium* (*Epidendroideae*), i geni di funzione B, C e D si co-esprimono nella colonna, confermando la natura chimerica di tale struttura (fusione degli organi riproduttivi maschili e femminili) (Skipper et al., 2006; Song et al., 2006; Xu et al., 2006). In *Dendrobium*, i geni di classe C e D sono espressi in tutti gli organi fiorali. In particolare, si osserva un aumento del livello di trascrizione dei geni di funzione D durante l'ultimo stadio dello sviluppo degli ovuli; tali risultati suggeriscono che i geni di classe C e D agiscono in concerto con altri geni per specificare l'identità di tutti gli organi fiorali, ed i geni di classe D hanno un ruolo importante nell'ultima fase della formazione degli ovuli (Skipper et al., 2006; Song et al., 2006; Xu et al., 2006).

L' "Orchid code"

Come descritto in precedenza, i fiori delle orchidee sono caratterizzati da un perianzio molto particolare costituito da tepali esterni, tepali interni laterali e da un tepalo interno mediano, il labello. Sulla base di studi condotti di recente, è stata sviluppata una teoria per spiegare l'ontogenesi degli organi fiorali che costituiscono il perianzio delle orchidee, nota come "orchid code", secondo cui gli eventi duplicativi seguiti da

diversificazione funzionale dei geni di classe B, in particolare dei geni *DEF-like*, hanno svolto un ruolo fondamentale nell'origine evolutiva delle peculiarità morfologiche che si osservano nel fiore delle orchidee (Hsu et al., 2002; Tsai et al., 2004; Xu et al., 2006; Kim et al., 2007; Mondragon-Palomino e Theißen, 2008).

Nelle *Orchidaceae*, i geni *DEF-like* formano quattro gruppi ben distinti: *PeMADS2-like* (clade 1), *OMADS3-like* (clade 2), *PeMADS3-like* (clade 3) e *PeMADS4-like* (clade 4); inoltre, i geni appartenenti ad uno specifico clade hanno lo stesso profilo di espressione: i geni del clade 1 e del clade 2 si esprimono in tutti i tepali, i geni del clade 3 solo nei tepali interni laterali e i geni del clade 4 solo nel labello (Figura 10) (Tsai et al., 2004; Xu et al., 2006; Kim et al., 2007). E' stato ipotizzato che la differenza morfologica tra tepali esterni ed interni è determinata dall'espressione differenziale dei geni del clade 3 e la differenza tra tepali interni e labello dall'espressione differenziale dei geni del clade 4 (Mondragon-Palomino e Theißen, 2008).

La teoria dell'"orchid code" è supportata dai profili di espressione dei geni di classe B osservati nelle orchidee selvatiche (*wild-type*) e nei mutanti omeotici; ad esempio, in *Habenaria radiata*, l'espressione eterotopica del gene del clade 3 *HrDEF* nei tepali esterni determina la trasformazione dei tepali esterni in organi simili ai tepali interni laterali (Kim et al., 2007); in *P. equestris*, l'espressione eterotopica del gene di clade 4 *PeMADS4* nei tepali interni laterali determina la formazione di strutture simili al labello piuttosto che di tepali interni laterali (Tsai et al., 2004). Recentemente, inoltre, la teoria dell'"orchid code" è stata utilizzata per spiegare l'esistenza in natura di orchidee mutanti piloriche e pseudopiloriche, nelle quali si riscontra un profilo di espressione alterato dei geni omeotici, in particolare dei geni *DEF-like* (Mondragon-Palomino et al., 2009).

Evoluzione dei geni *DEF-like* nelle orchidee

I geni *DEF-like* nelle orchidee hanno avuto origine da due eventi duplicativi avvenuti prima della separazione delle *Vanilloideae*:

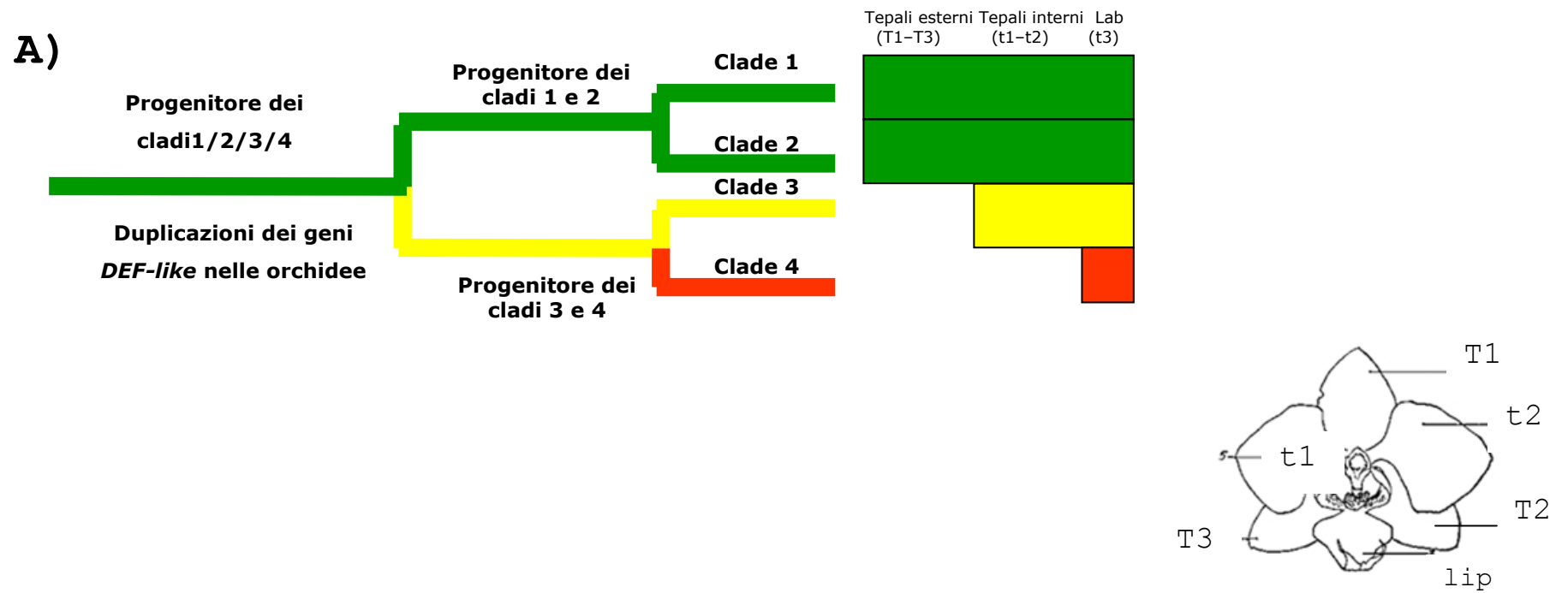


Figura 10 - Rappresentazione grafica della teoria dell'"orchid code" (A) e schematizzazione del fiore dell'orchidea *Phalaenopsis* (B). T1, T2 e T3 - tepali esterni; t1, t2 - tepali interni; lip - labello.

la prima duplicazione ha dato origine al progenitore dei geni dei cladi 1 e 2 e al progenitore dei geni dei cladi 3 e 4, la seconda duplicazione ha generato i geni dei cladi 1, 2, 3 e 4 (Mondragon-Palomino et al., 2008).

Recentemente, è stato dimostrato che, dopo le duplicazioni descritte, i geni *DEF-like* hanno subito una diversificazione funzionale (sub-neofunzionalizzazione) dovuta ad un rilassamento delle costrizioni selettive agenti sui domini K e C-terminale e sui loro promotori che ha determinato nuove interazioni proteina-proteina e nuovi profili di espressione, influenzando lo sviluppo fiorale delle orchidee e contribuendo alla loro diversificazione (Mondragon-Palomino et al., 2009).

L'evoluzione dei geni *DEF-like* nelle orchidee ha determinato, quindi, la nascita di un perianzio "modularizzato" costituito ossia da tepali capaci di evolvere in maniera semi-indipendente (Mondragon-Palomino et al., 2008).

Gli studi evolutivi condotti sui geni *DEF-like* hanno, inoltre, reso possibile ricostruire l'origine e l'evoluzione del perianzio delle orchidee: è stato ipotizzato che esso si sia evoluto da un progenitore caratterizzato da tepali indifferenziati (stato ancestrale); successivamente, come risultato della prima duplicazione, si è evoluto un perianzio con due tipi di tepali, esterni ed interni (stadio intermedio, morfologia osservata in *Apostasia*); la seconda duplicazione ha dato origine ad un fiore con una struttura zigomorfa ed un perianzio costituito da organi differenziati: tepali esterni, tepali interni laterali e labello (stato derivato) (Figura 11) (Mondragon-Palomino et al., 2008).

STATO ANCESTRALE

STATO INTERMEDIO

STATO DERIVATO

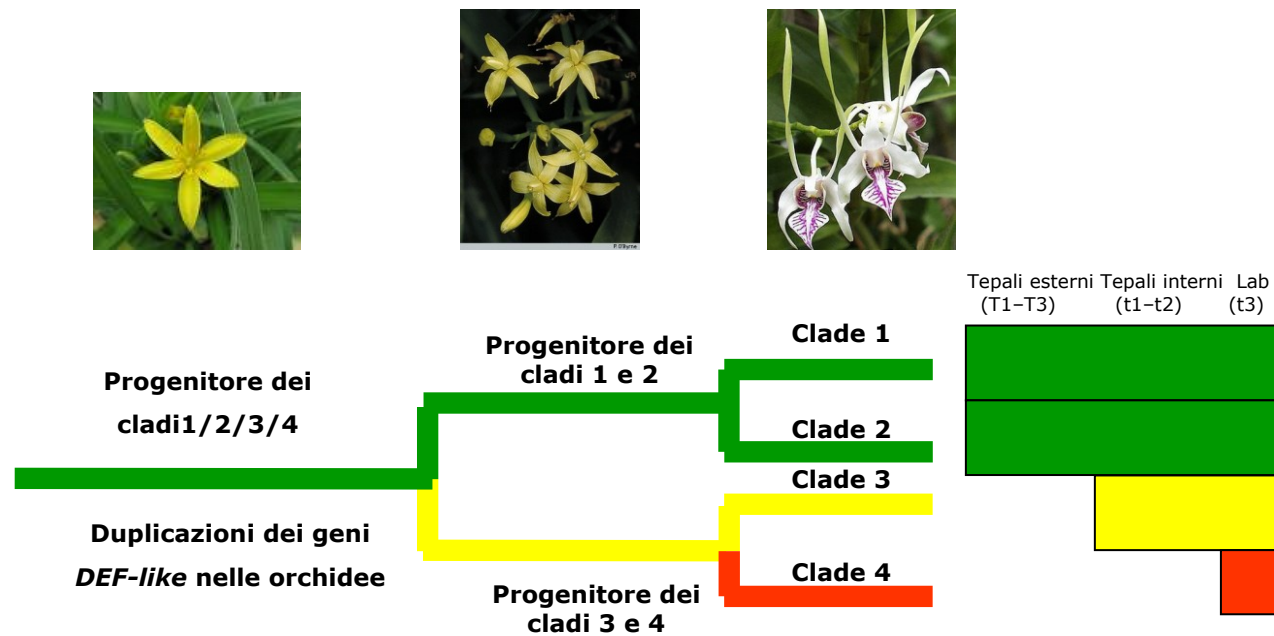


Figura 11 - Rappresentazione grafica dell'evoluzione dei geni *DEF-like* e l'origine del perianzio delle orchidee.

Scopo del lavoro

Il gruppo presso il quale ho svolto l'attività di ricerca di dottorato si interessa da alcuni anni di orchidee mediterranee, inizialmente per chiarire relazioni filogenetiche poco chiare tra alcuni *taxa*. L'analisi molecolare condotta su alcune specie di orchidee utilizzando i polimorfismi di lunghezza dei frammenti di restrizione (RFLP) del DNA plastidiale (Cozzolino et al., 1998) e le sequenze delle regioni spaziatrici trascritte (ITS, *Internal Transcribed Sequence*) del DNA ribosomale (Aceto et al., 1999) hanno evidenziato che, per alcune specie, l'evoluzione di alcuni caratteri fiorali non è in accordo con quella molecolare.

Ad esempio, *Orchis italica* e *Orchis anthropophora* (ex *Aceras antropophorum*), due specie che, analizzate sulla base della morfologia dei caratteri fiorali, erano classificate come appartenenti a generi differenti, sono risultate filogeneticamente molto vicine sulla base dell'analisi dei marcatori molecolari prima citati; infatti, da un punto di vista molecolare, *O. antropophora* non è soltanto un membro altamente derivato del genere *Orchis*, ma anche uno dei *taxa* più vicini ad *O. italica* (Figura 12). La conclusione di tale studio è che, in alcuni gruppi di orchidee, i cambiamenti morfologici possano essere stati la conseguenza di minime modificazioni verificatesi in geni regolativi implicati nello sviluppo florale, piuttosto che il risultato di un lungo processo selettivo (Aceto et al., 1999).

Per cercare di spiegare l'assenza di corrispondenza tra evoluzione molecolare e morfologica osservata per alcune specie di orchidee analizzate, il gruppo di ricerca presso il quale ho svolto il mio lavoro ha indirizzato lo studio verso geni regolativi la cui alterazione può determinare variazioni evidenti della morfologia del fiore e/o dell'infiorescenza. Questo aspetto è stato l'oggetto della mia tesi di dottorato.

In particolare, ho isolato ed analizzato la variabilità di alcuni geni MADS-box per comprendere l'evoluzione molecolare dello sviluppo florale nelle specie di orchidee mediterranee analizzate.

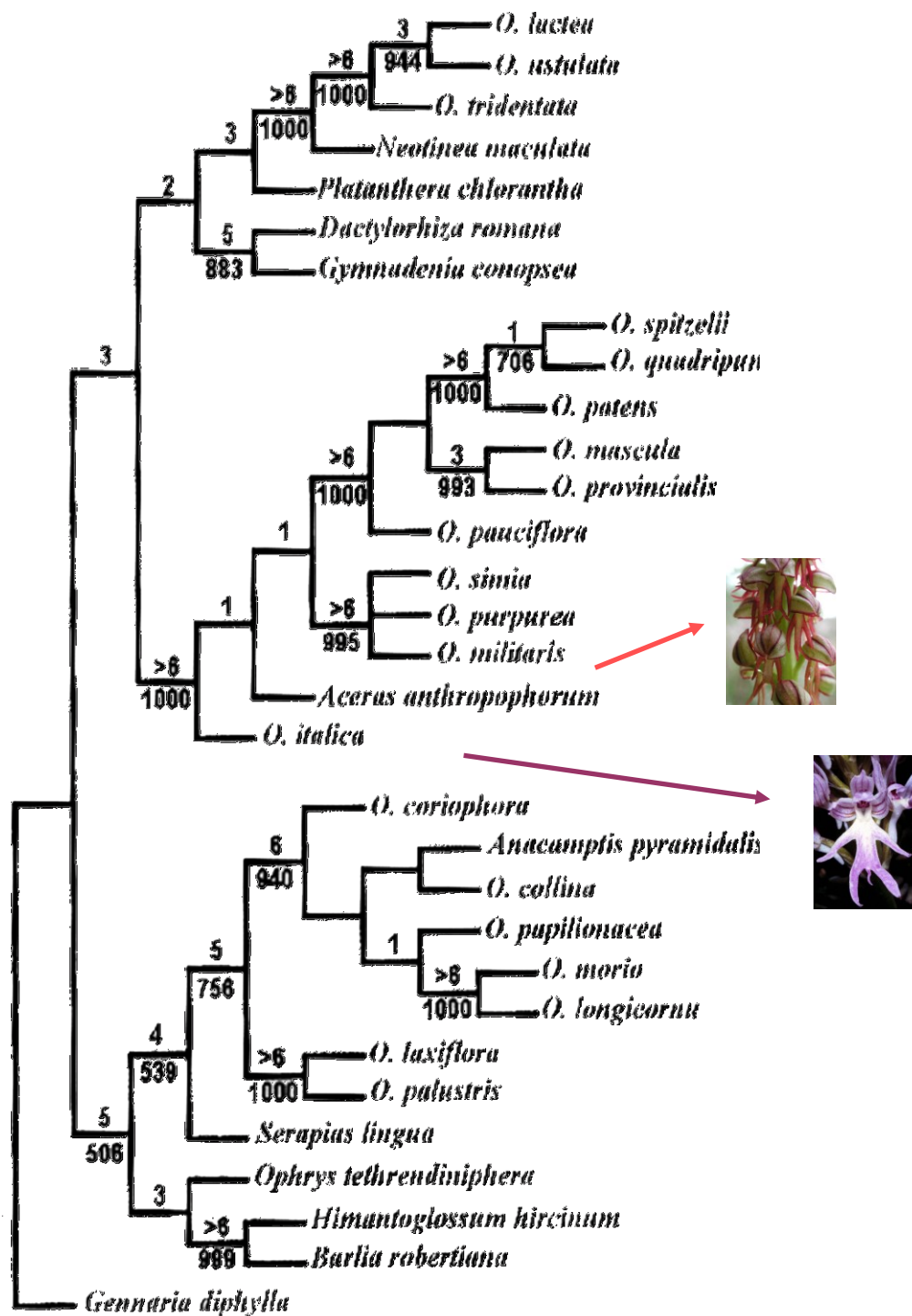


Figura 12 - Albero filogenetico di un gruppo di orchidee mediterranee ottenuto dall'analisi delle sequenze ITS 1 e ITS 2. (Da Aceto et al., Mol. Phyl. Evol. 1999). N.B.: la specie *Acercus anthropophorum* è oggi *Orchis anthropophora*; la specie *Orchis morio* è oggi *Anacamptis morio*; la specie *Orchis laxiflora* è oggi *Anacamptis laxiflora*.

Materiale

Gli esemplari utilizzati in questo lavoro (Tabella 1) sono presenti presso l'Orto Botanico di Napoli.

Per estrarre il DNA genomico, sono stati prelevati e congelati a -80°C frammenti di foglie (0,5 cm x 0,5 cm) da tutti i campioni in esame.

Per estrarre l'RNA totale, sono stati raccolti, immersi in RNA later e congelati a -80°C diversi tessuti di *O. italica*: radici, foglie, infiorescenze mature (in stadi differenti), infiorescenze senescenti, tepali esterni, tepali interni, labello, ovario, colonna.

Metodi

Allestimento e screening di una mini-libreria di cDNA di geni MADS-box in *O. italica*

L'RNA totale (1 μg), estratto da infiorescenza matura (100 mg) di *O. italica* utilizzando il *Trizol Reagent* (Invitrogen), è stato sottoposto alla reazione di RT-PCR (*RetroTranscribed-PCR*) utilizzando il *kit Advantage RT for PCR* (Roche), nelle condizioni consigliate dalla casa produttrice. Il primo filamento di cDNA è stato sintetizzato utilizzando un *primer* oligo-dT (5'-GAATTCGGATCC(T)₁₆-3') che, ibridando con la coda di poli-A degli mRNA, ha escluso tutti gli altri tipi di RNA dalla sintesi. Il filamento neo-sintetizzato è stato amplificato mediante PCR utilizzando il *primer* oligo-dT ed un *primer* degenerato (5'-GGGAGRGGGAAGATMGAGAT-3') corrispondente a una sequenza molto conservata della regione che codifica il dominio MADS-box. La miscela di reazione è stata allestita in un volume finale di 50 μl con le seguenti concentrazioni di reagenti:

Tampone di reazione (Fermentas)	1X
MgCl ₂ (Fermentas)	1,5 mM
dNTPs (Fermentas)	0,2 mM
<i>primer forward</i>	1 μM
<i>primer reverse</i>	1 μM
<i>Taq</i> polimerasi (Fermentas)	1,5

Nome delle specie	Regioni analizzate
<i>Anacamptis laxiflora</i> ALAX, da 1 a 8	<ul style="list-style-type: none"> • Regione genomica del locus <i>OrcPI</i> che si estende dal primo al sesto esone, dal settimo esone alla 3'-UTR. • Promotore del locus <i>OrcPI</i> • Regione genomica del locus <i>OrcPI2</i> che si estende dal primo al sesto esone.
<i>Anacamptis morio</i> AMOR, da 1 a 9	<ul style="list-style-type: none"> • Regione genomica del locus <i>OrcPI</i> che si estende dal primo al sesto esone. • Promotore del locus <i>OrcPI</i> • Regione genomica del locus <i>OrcPI2</i> che si estende dal primo al sesto esone.
<i>Anacamptis pyramidalis</i> APYR	<ul style="list-style-type: none"> • Regione genomica del locus <i>OrcPI</i> che si estende dal primo al sesto esone. • Promotore del locus <i>OrcPI</i> • Regione genomica del locus <i>OrcPI2</i> che si estende dal primo al sesto esone.
<i>Dactylorhiza romana</i> DROM	<ul style="list-style-type: none"> • Regione genomica del locus <i>OrcPI</i> che si estende dal primo al sesto esone. • Promotore del locus <i>OrcPI</i>
<i>Gennaria diphilla</i> GDIP	<ul style="list-style-type: none"> • Regione genomica del locus <i>OrcPI</i> che si estende dal primo al sesto esone. • Promotore del locus <i>OrcPI</i> • Regione genomica del locus <i>OrcPI2</i> che si estende dal primo al sesto esone.
<i>Himantoglossum robertianum</i> HROB	<ul style="list-style-type: none"> • Regione genomica del locus <i>OrcPI</i> che si estende dal primo al sesto esone. • Promotore del locus <i>OrcPI</i>
<i>Neotinea lactea</i> NLAC	<ul style="list-style-type: none"> • Regione genomica del locus <i>OrcPI</i> che si estende dal primo al sesto esone. • Promotore del locus <i>OrcPI</i> • Regione genomica del locus <i>OrcPI2</i> che si estende dal primo al sesto esone.
<i>Neotinea maculata</i> NMAC	<ul style="list-style-type: none"> • Regione genomica del locus <i>OrcPI</i> che si estende dal primo al sesto esone. • Promotore del locus <i>OrcPI</i> • Regione genomica del locus <i>OrcPI2</i> che si estende dal primo al sesto esone.
<i>Ophrys tenthrediniphora</i> OPTE	<ul style="list-style-type: none"> • Regione genomica del locus <i>OrcPI</i> che si estende dal primo al sesto esone. • Promotore del locus <i>OrcPI</i> • Regione genomica del locus <i>OrcPI2</i> che si estende dal primo al sesto esone.

<i>Orchis antropophora</i>	OANT	<ul style="list-style-type: none"> • Regione genomica del locus <i>OrcPI</i> che si estende dal primo al sesto esone. • Promotore del locus <i>OrcPI</i> • Regione genomica del locus <i>OrcPI2</i> che si estende dal primo al sesto esone.
<i>Orchis italica</i>	OITA	<ul style="list-style-type: none"> • cDNA dei geni <i>OrcAG</i>, <i>OrcDEF</i> e <i>OrcPI</i> • 5'-UTR dei loci <i>OrcPI</i> e <i>OrcPI2</i> • Regione genomica del locus <i>OrcPI</i> che si estende dal primo al sesto esone. • Promotore del locus <i>OrcPI</i> • Regione genomica del locus <i>OrcPI2</i> che si estende dal primo al sesto esone.
<i>Serapias lingua</i>	SLIN	<ul style="list-style-type: none"> • Regione genomica del locus <i>OrcPI</i> che si estende dal primo al sesto esone. • Promotore del locus <i>OrcPI</i>
<i>Orchis mascula</i>	OMAS	<ul style="list-style-type: none"> • Regione gnomica del locus <i>OrcPI</i> che si estende dal primo al sesto esone. • Promotore del locus <i>OrcPI</i> • Regione genomica del locus <i>OrcPI2</i> che si estende dal primo al sesto esone.

Tabella 1 - Elenco delle specie e delle regioni dei geni analizzate nel presente lavoro.

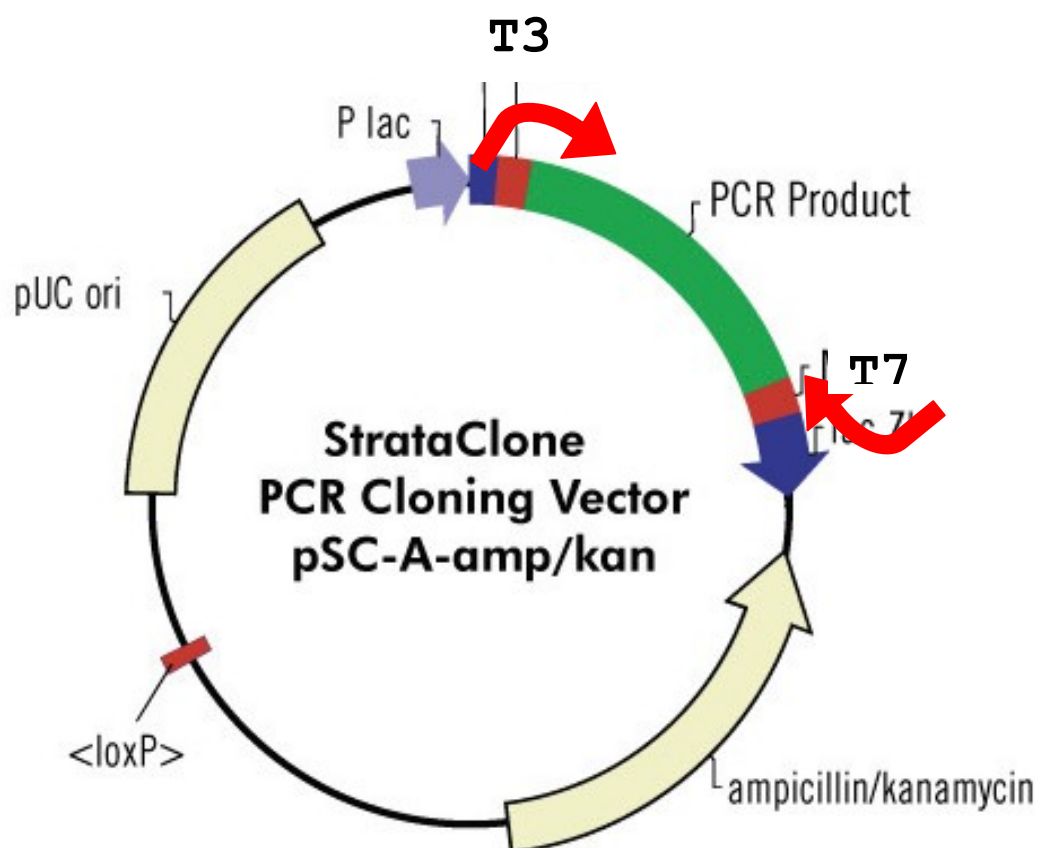
Il programma termico, condotto nell'apparecchio *DNA MiniThermal Cyclers* (MJ Research), è il seguente:

94°C	3 min	
94°C	30 sec	} 30 cicli
50°C	30 sec	
72°C	1 min	
72°C	7 min	

Il prodotto di amplificazione è stato clonato nel vettore pSCA (*StrataClone PCR Cloning Kit*, Stratagene) e i cloni ricombinanti sono stati sequenziati con i primer "universali" T7 e T3, specifici della maggior parte dei vettori di clonaggio, che ibridano a monte e a valle del sito di policlonaggio (Figura 13). Venti differenti cloni positivi per la presenza di un inserto sono stati selezionati per il sequenziamento. Le reazioni di sequenza sono state condotte utilizzando il kit "*Big Dye Terminator v1.1*" (Applied Biosystems) e analizzate al Sequenziatore Automatico Capillare ABI 310 (Applied Biosystems).

Le sequenze ottenute sono state prima controllate manualmente e poi allineate utilizzando il software *BioEdit* (Hall, 1999). L'analisi BLAST (*Basic Local Analysis Sequence Tool*) ha consentito di confrontare le sequenze ottenute con quelle presenti in banca dati (GenBank) identificando, durante questo *screening*, otto differenti cDNA MADS-box in *O. italica*: due AG-like (chiamati *OrcAG1* e *OrcAG2*), quattro DEF-like (chiamati *OrcDEF1*, *OrcDEF2*, *OrcDEF3* e *OrcDEF4*), due GLO-like (chiamati *OrcPI* e *OrcPI2*).

Le sequenze di cDNA sono state utilizzate per disegnare primer specifici per amplificare diverse regioni genomiche dei loci MADS-box identificati, con particolare attenzione ai loci GLO-like (*OrcPI* e *OrcPI2*). Tale scelta, come sarà meglio discusso nei paragrafi successivi, è dovuta al fatto che nelle orchidee il locus GLO-like è presente in singola copia (Mondragon-Palomino et al., 2009), con l'unica eccezione di *Habenaria radiata* (Kim et al., 2007), specie appartenente, come *O. italica*, alla sottofamiglia *Orchidoideae*.



T3 5' -GAAATTAACCCTCACTAAAG-3'

T7 5' -TAATACGACTCACTATAGGG-3'

Figura 13 - Rappresentazione grafica del vettore pSC-A e sequenze dei primer T3 e T7 (Strataclone).

Identificazione delle regioni 5'UTR dei geni *OrcPI* e *OrcPI2* in *O. italica*

Per isolare la regione presente all'estremità 5' del cDNA dei geni *OrcPI* e *OrcPI2*, 1 µg di RNA totale estratto da infiorescenze di *O. italica* è stato sottoposto alla reazione di 5' RACE (*Rapid Amplification of cDNA End*) utilizzando il kit "5'/3' RACE, 2nd generation" (Roche).

L'estremità 5'UTR dei geni è stata identificata tramite due successive reazioni di PCR, utilizzando *primer reverse* gene-specifici, disegnati sulla base della sequenza dei cDNA di *OrcPI* e *OrcPI2*, nelle condizioni consigliate dalla casa produttrice del kit (Figura 14). Le sequenze dei *primer* sono riportate nella Tabella 2, la loro posizione è indicata nella Figura 15.

Il prodotto della seconda amplificazione è stato clonato nel vettore pSCA (*StrataClone PCR Cloning Kit*, Stratagene) e i cloni ricombinanti sono stati sequenziati con i *primer* T3 e T7.

Identificazione del promotore del gene *OrcPI* in *O. italica*

Per identificare il promotore del gene *OrcPI* è stata utilizzata la tecnica *Thermal Asymmetric Interlaced* PCR (TAIL-PCR), ideata come strategia per isolare segmenti di DNA adiacenti a regioni con sequenza nota (Liu e Whitter, 1995), quindi anche di promotori (Terauchi e Kahl., 2000; Li e Gray, 2005).

La reazione di TAIL-PCR prevede l'utilizzo combinato di *primer* specifici del gene e di *primer* con sequenza arbitraria, degenerati o meno. Le tecniche di PCR che presentano tale caratteristica sono note come "PCR semi-specifiche"; i prodotti amplificati tramite PCR semi-specifica possono essere di triplice natura: ampliconi ottenuti da entrambi i *primer* (tipo I), ampliconi ottenuti dal solo *primer* specifico (tipo II) e ampliconi ottenuti dal solo *primer* con sequenza arbitraria (tipo III). L'amplificazione preferenziale della regione bersaglio (prodotto di tipo I) è controllata termicamente alternando cicli di PCR ad alta stringenza, che favoriscono l'appaiamento del *primer* specifico,

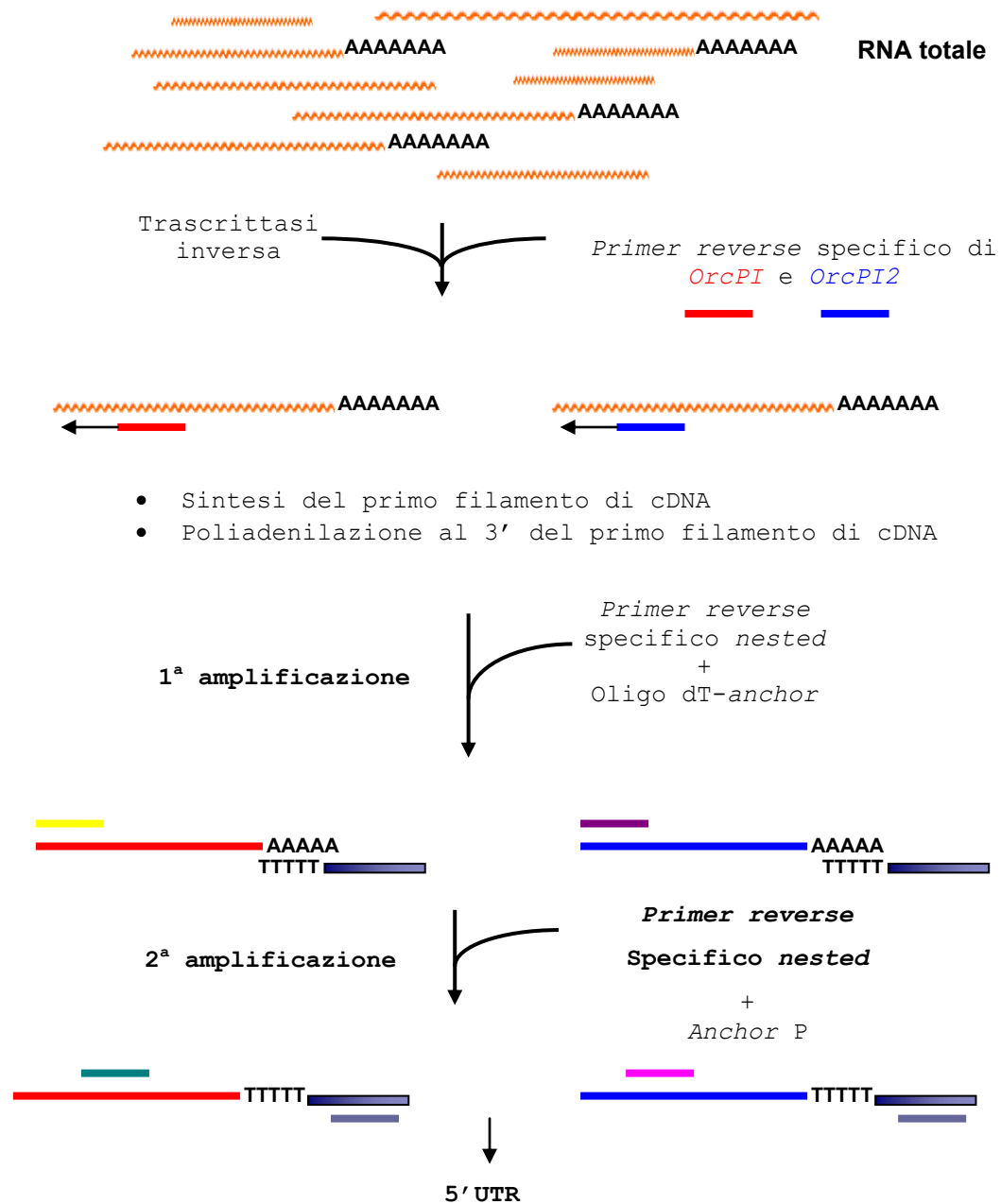


Figura 14 - Rappresentazione grafica della 5'RACE condotta per isolare la 5'-UTR dei geni *GLO-like* di *O. italica*.

Nome <i>primer</i> (<i>OrcPI2</i>)	Sequenza (5' → 3')
GLOR1	5'-GAAGGCGAAGTCAGCAGATT-3'
GLOR2	5'-GTCAGACAGATATATCTAACT-3'
GLOR3	5'-TATCGCGAACACCAGCCAGA-3'
Nome <i>primer</i> (<i>OrcPI</i>)	Sequenza (5' → 3')
PI1	5'-CAGTTCTCTCATGCTGCCTTC-3'
PI2	5'-CTGAGCTCGATCTGCATGT-3'
PI3	5'-GCTGCAGTATTCGGACAA-3'

Tabella 2 - Sequenze dei *primer* utilizzati nella 5'-RACE.

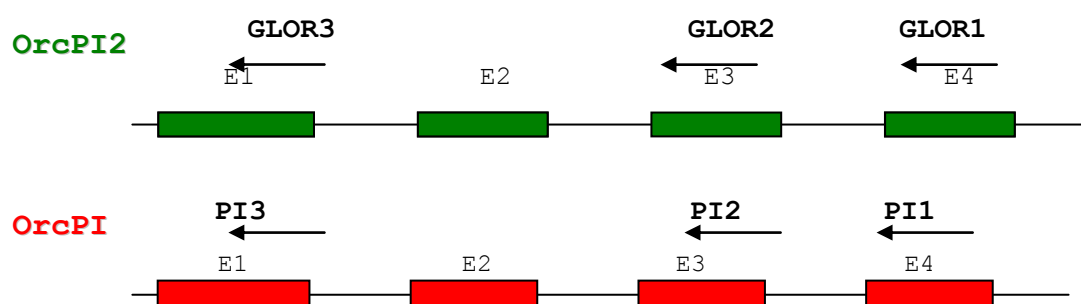


Figura 15 - Posizione dei *primer* utilizzati nella 5'-RACE.

con cicli a bassa stringenza che facilitano l'appaiamento del *primer* con sequenza arbitraria. Dopo la prima reazione di amplificazione, i prodotti aspecifici sono diluiti attraverso una seconda reazione di PCR usando un *primer* interno (*nested*) specifico (Liu e Whitter, 1995) (Figura 16).

Per isolare il promotore del gene *OrcPI* è stato necessario apportare diverse modifiche al protocollo iniziale di TAIL-PCR (Liu e Whitter, 1995) riguardanti vari parametri: concentrazione di DNA genomico, concentrazione dei *primer*, concentrazione di $MgCl_2$, volume finale di reazione, condizioni termiche.

Circa 200 ng di DNA genomico di *O. italica* sono stati utilizzati per condurre la prima reazione di amplificazione (TAIL-1) in un volume finale di 20 μ l con le seguenti concentrazioni di reagenti:

Tampone di reazione (Fermentas)	1X
$MgCl_2$ (Fermentas)	2 mM
dNTPs (Fermentas)	0,2 mM
<i>primer forward random</i>	2 μ M
<i>primer reverse OrcPI-specifico</i>	0,2 μ M
<i>Taq</i> polimerasi (Fermentas)	2 U

I prodotti della reazione TAIL-1 sono stati sottoposti ad una seconda amplificazione (TAIL-2) allestita con le stesse concentrazioni di reagenti della prima amplificazione in un volume finale di 30 μ l, utilizzando 2 μ l di una diluizione 1:200 (v/v) della TAIL-1. Le combinazioni dei *primer* usate nella TAIL-1 e nella TAIL-2 sono riportate nella Tabella 3; il programma termico utilizzato è descritto nella Tabella 4.

Le reazioni della TAIL-2 sono state analizzate per elettroforesi su un gel d'agarosio 1,5%. I frammenti di amplificazione sono stati prelevati dal gel, eluiti e amplificati tramite PCR per verificare la loro specificità, usando *primer OrcPI-specifici* che amplificano un frammento di 156 bp corrispondente al primo esone del gene. Le reazioni di amplificazione sono state allestite in un volume finale di 50 μ l con le seguenti concentrazioni di reagenti:

Tampone di reazione (Fermentas)	1X
$MgCl_2$ (Fermentas)	1,5 mM
dNTPs (Fermentas)	0,2 mM
<i>primer forward</i>	1 μ M

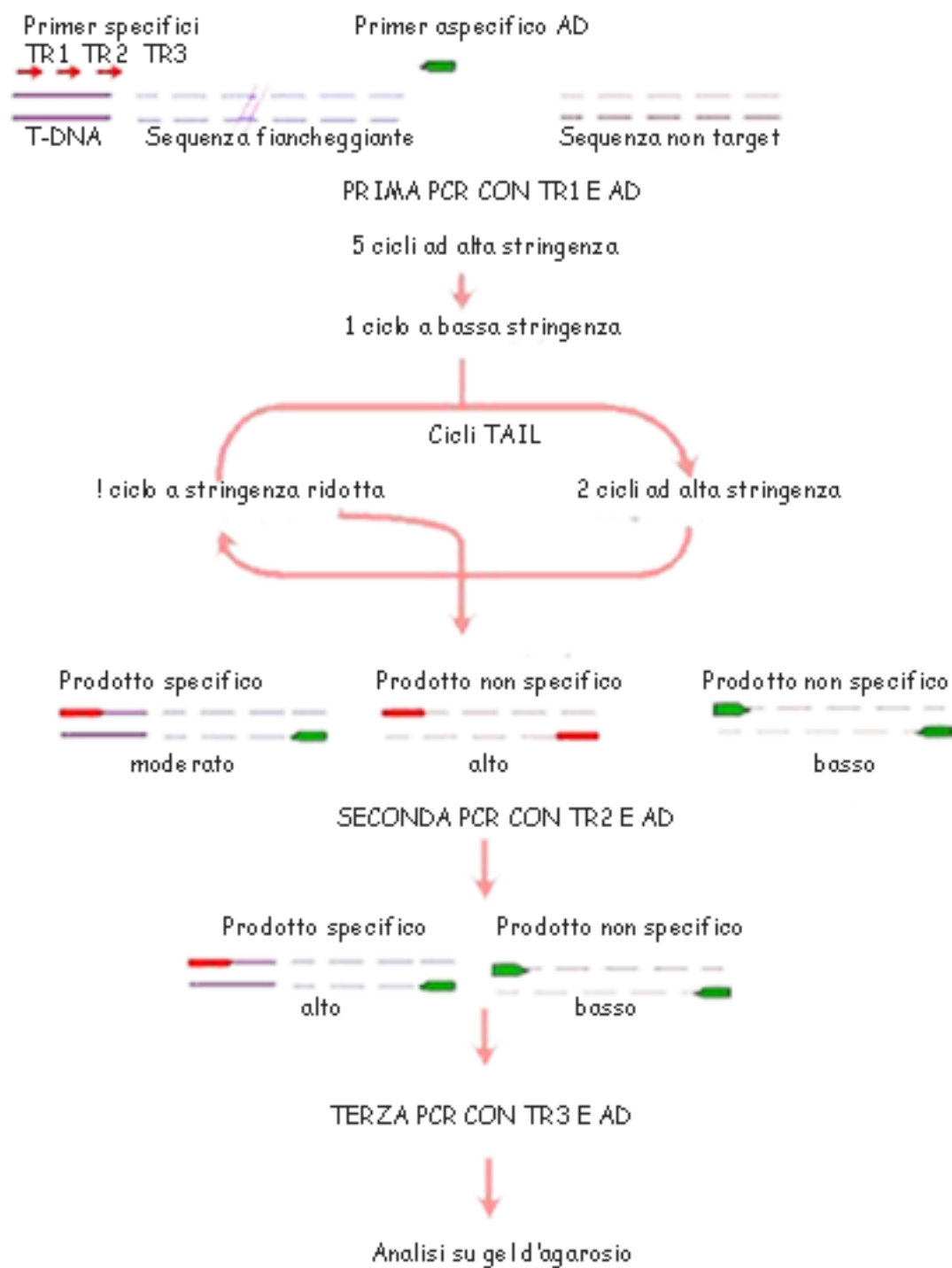


Figura 16 - Rappresentazione grafica della TAIL-PCR.

Nome	Sequenza (5'–3')	Tipo
R1	ggtgcgggaa	<i>Forward, random</i>
R2	gtttcgctcc	<i>Forward, random</i>
R3	gtagaccgt	<i>Forward, random</i>
R4	aagagcccgt	<i>Forward, random</i>
R5	aacgcgcaac	<i>Forward, random</i>
R6	cccgctagca	<i>Forward, random</i>
R7	cctcaagatt	<i>Forward, random</i>
R8	wgtgnagwancanaga	<i>Forward, random</i>
LSER I TAIL-PCR	gtggtagaggggctgca	<i>Reverse, OrcPI-specifico</i>
VIFR II TAIL-PCR	gcctaggctggagaagatgac	<i>Reverse, OrcPI-specifico</i>
PIF Controllo	atggggcggggaaatacggag	<i>Forward, OrcPI-specifico</i>
GIIR Controllo	ctccctcgctttctttatgattcc	<i>Reverse, OrcPI-specifico</i>

Tabella 3 - Sequenze dei primer utilizzati nella TAIL-PCR.

Reazione	Numero dei cicli	Condizioni termiche
TAIL1	1	95°C-2 min
	5	94°C-30 s, 60°C-1 min, 72°C-2.5 min
	1	94°C-30 s, 25°C-3 min, 72°C-2.5 min (ramping 25-72 0.3°C/s)
	15	94°C-10 s, 60°C-1 min, 72°C-2.5 min 94°C-10 s, 60°C-1 min, 72°C-2.5 min 94°C-10 s, 44°C-1 min, 72°C-2.5 min
	1	72°C-5 min, 14°C hold
TAIL 2	1	95°C-2 min
	12	94°C-10 s, 64°C-1 min, 72°C-2.5 min 94°C-10 s, 64°C-1 min, 72°C-2.5 min 94°C-10 s, 44°C-1 min, 72°C-2.5 min
	1	72°C-5 min, 14°C hold

Tabella 4 - Programma termico della TAIL-PCR.

<i>primer reverse</i>	1 μ M
<i>Taq</i> polimerasi (Fermentas)	1,5 U

Le condizioni termiche utilizzate sono le seguenti:

94°C	3 min	
94°C	30 sec	} 35 cicli
60°C	30 sec	
72°C	30 sec	
72°C	7 min	

Le sequenze dei *primer* sono riportate nella Tabella 3.

I prodotti della reazione TAIL-2 positivi all'amplificazione del primo esone del gene *OrcPI* sono stati clonati (*StrataClone PCR kit*, Stratagene) e sequenziati usando i *primer* T7 e T3.

Analisi funzionale del promotore del gene *OrcPI*

Per verificare che la regione isolata a monte del gene *OrcPI* fosse il promotore di tale gene è stato necessario allestire saggi di espressione *in vivo* basati sulla strategia della fusione genica: l'intera regione isolata e frammenti progressivamente deleti all'estremità 5' sono stati clonati nel vettore di espressione pBI121 (Clontech), a monte del gene *reporter GUS*.

Il gene *GUS* è usato come gene *reporter* per diversi motivi: la facilità con la quale l'attività della proteina codificata dal gene (β -glucuronidasi) può essere rilevata, l'assenza di un'attività endogena dell'enzima (basso rumore di fondo) e la notevole sensibilità del saggio (Jefferson et al., 1987).

Il gene *GUS* codifica la β -glucuronidasi, enzima in grado di idrolizzare i legami β -glicosidici. E' possibile quindi studiare il ruolo funzionale delle sequenze promotrici clonate a monte del gene *GUS* analizzando l'attività enzimatica della β -glucuronidasi tramite un saggio istochimico. In presenza di un substrato artificiale, l'X-gluc (5-bromo-4-cloro-3-indol- β -glucuronide), la reazione di idrolisi proteica produce, attraverso un intermedio instabile, un precipitato blu insolubile, il dicloro-dibromo indingo, che immediatamente precipita. Questo permette di

determinare se la regione clonata a monte del gene *GUS* sia in grado di promuoverne la trascrizione (funzione di promotore).

Il plasmide pBI121 è un vettore binario di 15 Kb. Contiene, all'interno del T-DNA, fra le sequenze *Right Border* (RB) e *Left Border* (LB), il gene *NPT* che conferisce la resistenza all'antibiotico kanamicina, sotto il controllo del promotore pNOS; un sito di policlonaggio (contenente siti unici per gli enzimi di restrizione *HindIII*, *SalI*, *XbaI*, *BamHI*, *SmaI*) inserito all'interno della sequenza del promotore costitutivo 35S del virus del mosaico del tabacco; a valle del sito di policlonaggio è presente il gene *GUS* (Figura 17). Il vettore pBI121 è stato utilizzato per ottenere la fusione trascrizionale e traduzionale tra la sequenza isolata all'estremità 5' del gene *OrcPI* e il gene reporter *GUS*.

L'intero frammento di DNA genomico (1324 bp) isolato a monte della regione codificante del gene *OrcPI* e una serie di frammenti progressivamente deleti all'estremità 5' (di lunghezza 854, 577 e 356 bp) sono stati amplificati tramite PCR (Figura 18). Le sequenze dei *primer* utilizzati nelle reazioni di amplificazione sono riportate nella Tabella 5. Le reazioni di amplificazione sono state allestite utilizzando 100 ng di DNA genomico di *O. italica* in un volume finale di 50 µl con le seguenti concentrazioni di reagenti:

Tampone di reazione (Fermentas)	1X
MgCl ₂ (Fermentas)	1,5 mM
dNTPs (Fermentas)	0,2 mM
<i>primer forward</i>	1 µM
<i>primer reverse</i>	1 µM
<i>Taq</i> polimerasi (Fermentas)	1,5 U

Le condizioni termiche utilizzate sono le seguenti:

94°C	3 min	} 35 cicli
94°C	45 sec	
54°C	45 sec	
72°C	1 min	
72°C	7 min	

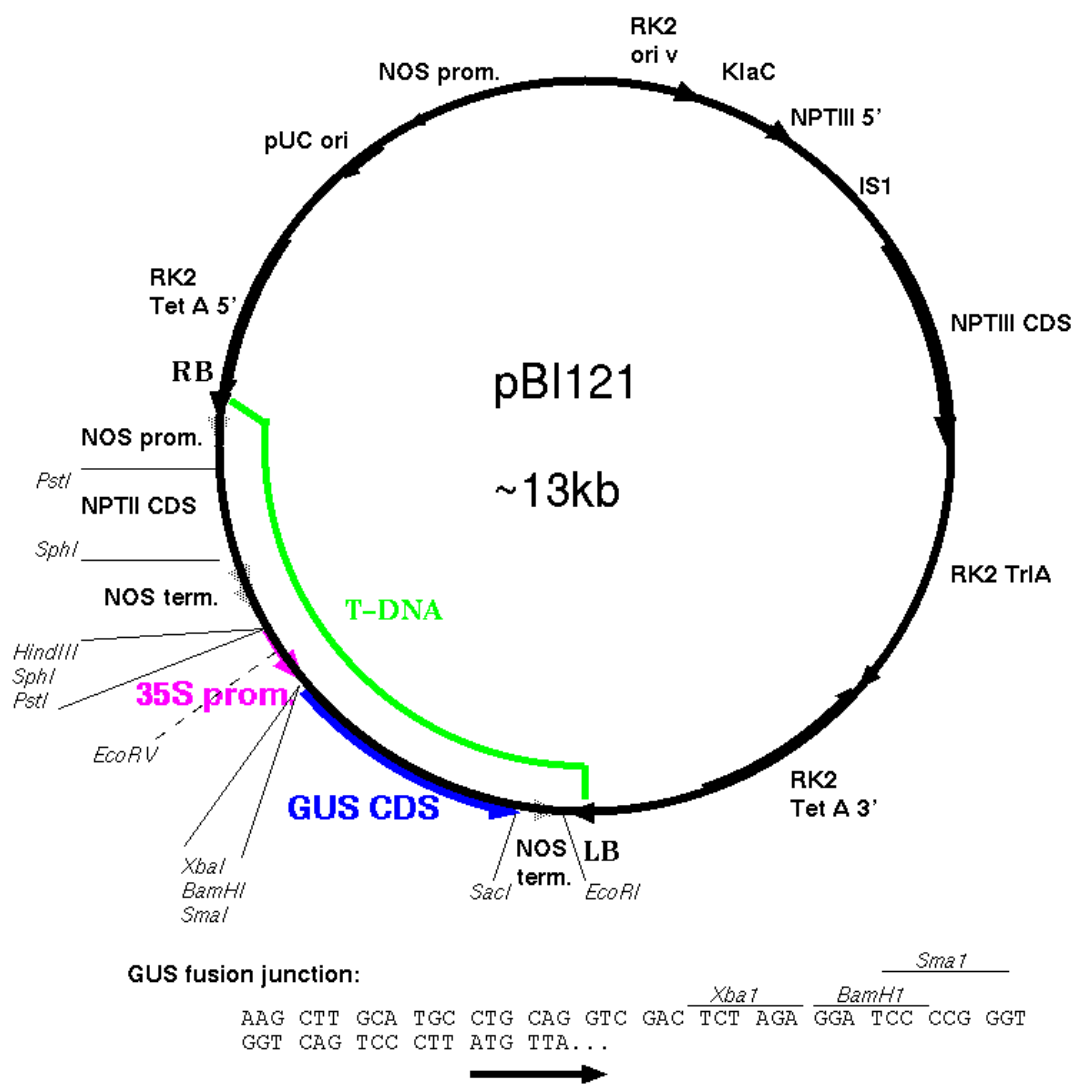


Figura 17 - Rappresentazione grafica del vettore binario pBI121.

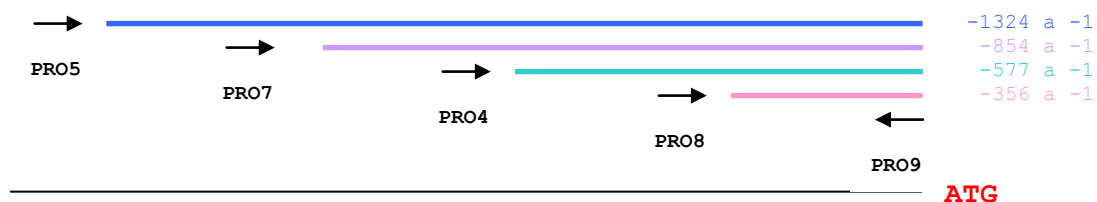


Figura 18 - Posizione dei *primer* utilizzati per amplificare i frammenti di *OrcPI* deleti all'estremità 5'.

Nome	Sequenza (5'-3')	Tipo
PRO5 (-1324)	ggagtaaacatgacaatagct	<i>Forward</i> , promotore-specifico
PRO7 (-854)	gtgaaaactatattgcgtcgt	<i>Forward</i> , promotore-specifico
PRO4 (-577)	tgatgcatcttgtgtctcaat	<i>Forward</i> , promotore-specifico
PRO8 (-356)	agagttgctctaagtaccgtt	<i>Forward</i> , promotore-specifico
PRO9 (-1)	ctcggaataacaggaaaaaat	<i>Reverse</i> , promotore-specifico

Tabella 5 - Sequenze dei *primer* utilizzati per amplificare i frammenti deleti all'estremità 5' del promotore di *OrcPI*.

I prodotti di amplificazione sono stati clonati nel vettore pSCA (Stratagene) e successivamente escissi tramite digestione enzimatica, utilizzando le endonucleasi di restrizione *HindIII* e *XbaI* (per i frammenti di 356, 577 e 854 bp) ed *EcoRV* e *SmaI* (per il frammento di 1324 bp). I frammenti ottenuti sono stati sub-clonati nel vettore di espressione pBI121, in maniera direzionale rispetto al gene *GUS*. L'inserzione nella direzione corretta è stata verificata tramite sequenziamento.

I costrutti chimerici generati, P1324OrcPI (contenete l'intero frammento isolato a monte del gene *OrcPI*), ΔP854OrcPI, ΔP577OrcPI e ΔP356OrcPI (contenenti i frammenti progressivamente deleti all'estremità 5') sono stati utilizzati per trasformare cellule del batterio *Agrobacterium tumefaciens* (ceppo AGL0). Oltre ai costrutti chimerici, sono state condotte trasformazioni di *A. tumefaciens* anche con il vettore pBI121 contenente il promotore costitutivo *CaMV 35S* a monte del gene *GUS* (controllo positivo) e con il vettore pBI121 *promoter-less*, privo del promotore costitutivo (controllo negativo) (Figura 19).

Per studiare l'attività funzionale delle regioni clonate, petali di *Rosa hybrida* sono stati agro-infiltrati con i ceppi trasformati di *A. tumefaciens*, immersi in 100 µl di *GUS assay solution* (1 mM X-gluc, 100mM fosfato di sodio pH 7,0, 0,1% Triton) e incubati a 26°C per 2 e 5 giorni. I tessuti sono stati poi immersi in etanolo 70% e fotografati.

Caratterizzazione strutturale del locus *OrcPI* in *O. italica* e in altre specie di Orchidoideae

Promotore

La sequenza del promotore del gene *OrcPI* di *O. italica* è stata utilizzata per disegnare una serie di *primer* per amplificare tale regione in diverse specie di orchidee mediterranee (Tabella 1). L'amplificazione di tale regione nelle varie specie analizzate è stata condotta tramite due reazioni separate: una per amplificare le prime 640 bp e l'altra per amplificare la restante parte del promotore (Figura 20); per ciascuna regione, inoltre, sono state effettuate due successive reazioni di amplificazione, utilizzando

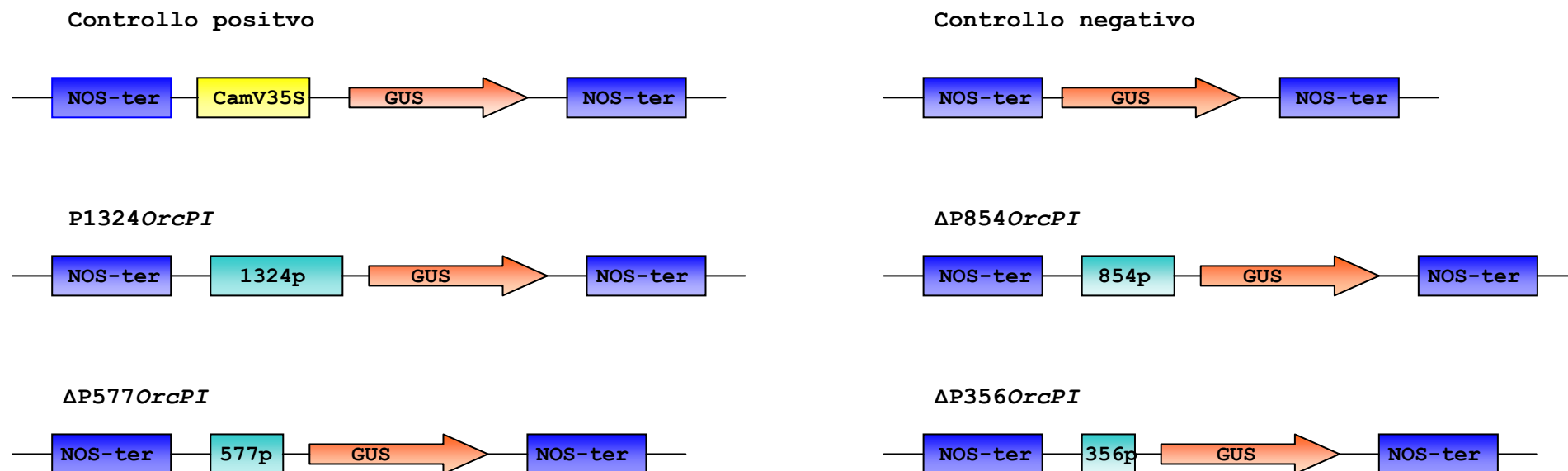
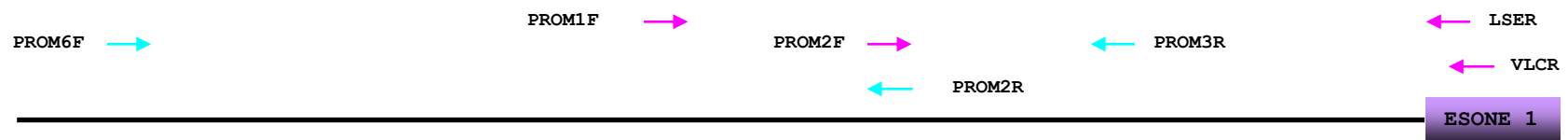


Figura 19 - Rappresentazione grafica dei costrutti chimerici del promotore di *OrcPI*.



PROM6F/PROM3R (I PCR)

PROM1F/VL5R (I PCR)

PROM6F/PROM2R (II PCR -798 bp promotore)

PROM2F/L5ER (II PCR - 640 bp promotore)

Figura 20 - Posizione dei *primer* utilizzati per amplificare il promotore di *OrcPI*.

nella seconda reazione di amplificazione 1 μ l di una diluizione 1:20 (v/v) della prima reazione.

Il DNA genomico (200 ng) estratto dalle specie in esame è stato amplificato tramite PCR allestendo reazioni con volume finale di 50 μ l con le seguenti concentrazioni di reagenti (nella seconda amplificazione sono state utilizzate le stesse condizioni):

Tampone di reazione (Fermentas)	1X
MgCl ₂ (Fermentas)	1,5 mM
dNTPs (Fermentas)	0,2 mM
<i>primer forward</i>	1 μ M
<i>primer reverse</i>	1 μ M
Taq polimerasi (Fermentas)	1,5 U

Le condizioni termiche utilizzate per amplificare il frammento iniziale di 670 bp sono le seguenti:

94°C	3 min	
94°C	45 sec	
56°C	45 sec	35 cicli
72°C	1 min	
72°C	7 min	

L'amplificazione del frammento di 798 bp è stata condotta con un analogo programma termico, con *annealing* a 60°C.

Le sequenze dei *primer* sono elencate nella Tabella 6. I prodotti amplificati sono stati clonati (*StrataClone PCR kit*, Stratagene) e sequenziati usando i primer T7 e T3.

Esoni e introni

In *O. italica* e nelle altre specie di orchidee in esame è stata amplificata la regione che si estende dal primo al sesto esone e quella che si estende dal settimo esone alla 3'UTR utilizzando *primer* disegnati sulla sequenza di cDNA del gene *OrcPI* di *O. italica* (Figura 21). Le sequenze dei *primer* sono elencate nella Tabella 7.

Nome primer	Sequenza (5'-3')
VLCR	GACCTGTGCGTCACAGAGCAC
LSEr	GCTGCAGTATTCGGACAACCTT
PROM3R	CACACGCGGGTAGACGCAGGC
PROM2R	GCCGTGGAACCCAACTCCACG
PROM2F	CGTGGAGTTGGGTTCCACGGC
PROM1F	TGGAGGCGTCGACGAGGAGCT
PROM6F	GGAGTAAACATGACAATAGCT

Tabella 6 - Sequenze dei *primer* utilizzati per amplificare il promotore di *OrcPI*.

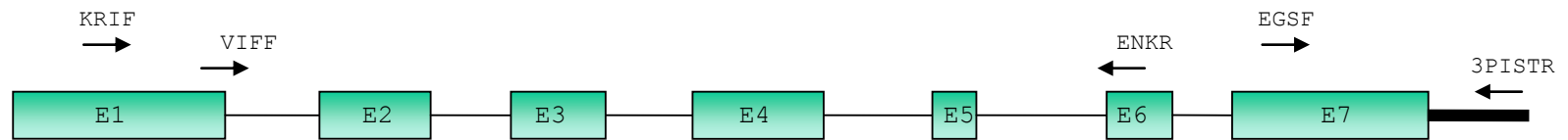


Figura 21 - Posizione dei *primer* utilizzati per amplificare la regione che si estende dal primo al sesto esone e dal settimo esone alla 3'-UTR del locus *OrcPI*.

Nome <i>primer</i>	Sequenza (5'-3')
KRIF	CAAGCGCATCGAGAACTCGACG
VIFF	GTCATCTTCTCCAGCCTAGGC
ENKR	ATATTTGAGCCTTTTATTTTC
EGSF	GAAGGCAGCATGAGAGAACTG
3PISTR	CGTTAAGAGAACTTTGGAGA

Tabella 7 - Sequenze dei *primer* utilizzati per amplificare esoni, introni e 3'-UTR del locus *OrcPI*.

Le miscele di reazione sono state allestite in un volume finale di 50 µl utilizzando 100 ng di DNA genomico e le seguenti concentrazioni di reagenti:

Tampone di reazione (Fermentas)	1X
MgCl ₂ (Fermentas)	1,5 mM
dNTPs (Fermentas)	0,2 mM
<i>primer forward</i>	1 µM
<i>primer reverse</i>	1 µM
Taq polimerasi (Fermentas)	1,5 U

Le reazioni sono state condotte con il seguente programma termico:

94°C	3 min	
94°C	30 sec	} 35 cicli
50°C	30 sec	
72°C	1 min	
72°C	7 min	

Per amplificare la regione che si estende dall'esone 1 all'esone 6, è stata condotta una seconda reazione di PCR nelle stesse condizioni descritte precedentemente e utilizzando come stampo 1 µl della prima reazione di amplificazione.

I frammenti amplificati sono stati clonati utilizzando il *kit* "StrataClone PCR" (Stratagene) e i cloni ricombinanti sono stati sequenziati (4 cloni differenti per ogni reazione di clonaggio) con i primer T3 e T7.

Analisi dell'espressione del gene *OrcPI* in *O. italica*

L'RNA totale è stato estratto (*Trizol Reagent*, Invitrogen) da foglie, radici, infiorescenze in 3 diversi stadi di sviluppo (1 cm lunghezza x 0,5 cm altezza; 1,5 cm x 0,8 cm; 2,5 cm x 1,4 cm; definiti come stadio 1, 2, e 3, rispettivamente), tepali esterni, tepali interni, labello, colonna, ovario, infiorescenza matura (6 giorni dopo la fioritura) e senescente (12 giorni dopo la fioritura) di *O. italica*.

Il primo filamento di cDNA sintetizzato a partire da 1 µg di RNA

totale (*Advantage RT for PCR kit*, Invitrogen) è stato amplificato tramite PCR utilizzando i *primer* specifici GLTF e PIR (Tabella 8), disegnati sulla regione del gene *OrcPI* codificante il dominio C-terminale (Figura 22).

Le miscele di reazione sono state allestite in un volume finale di 25 µl utilizzando 1 µl di cDNA con le seguenti concentrazioni di reagenti:

Tampone di reazione (Fermentas)	1X
MgCl ₂ (Fermentas)	1,5 mM
dNTPs (Fermentas)	0,2 mM
<i>primer forward</i>	1 µM
<i>primer reverse</i>	1 µM
<i>Taq</i> polimerasi (Fermentas)	1,5 U

Le reazioni sono state condotte con il seguente programma termico:

94°C	3 min	
94°C	30 sec	} 30 cicli
60°C	30 sec	
72°C	30 min	
72°C	7 min	

Per il cDNA di ovario sono state condotte altre due reazioni di amplificazione aumentando la quantità di cDNA stampo (2 µl e 4 µl). Come controllo, è stato amplificato un frammento di cDNA del gene che codifica l'actina utilizzando la coppia di *primer* ActinF e ActinR (Tabella 8).

Le reazioni di amplificazione sono state analizzate per elettroforesi su gel di agarosio 1,5%.

Caratterizzazione strutturale del locus *OrcPI2* in *O. italica* e in altre specie di *Orchidoideae*

In *O. italica* e nelle specie di orchidee elencate nella Tabella 1 è stata amplificata la regione che si estende dall'esone 1 all'esone 6 del gene *OrcPI2* tramite due successive reazioni di PCR utilizzando *primer* gene-specifici disegnati sulla sequenza di cDNA

Nome <i>primer</i>	Sequenze (5'-3')
GLTF	GGTTTGACTAGTGTTTCGGGAG
PIR	GTTGTTTCCGTGCAAGTTGGGCTG
ActinF	TGGAACTGCCAAGACG
ActinR	GCAGCGAAGATTCAAAA

Tabella 8 - Sequenza dei *primer* utilizzati nell'analisi di espressione del gene *OrcPI*.

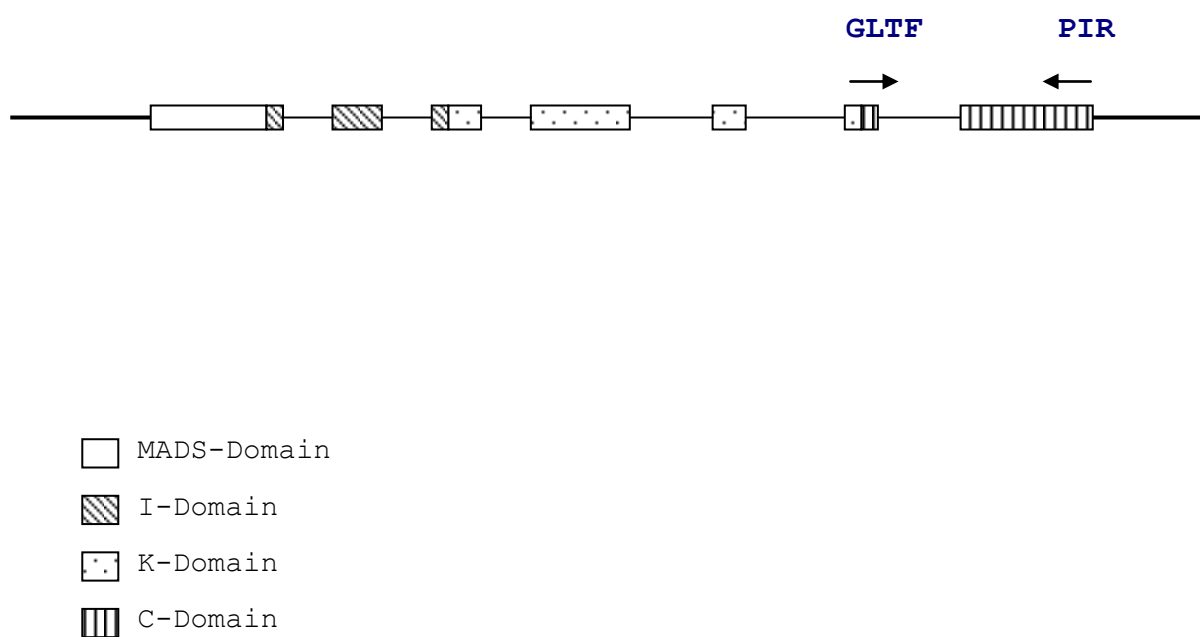


Figura 22 - Posizione dei *primer* utilizzati nell'analisi del profilo di espressione del gene *OrcPI*.

del gene *OrcPI2* di *O. italica* (Figura 23). Le sequenze dei *primer* sono elencate nella Tabella 9.

Le miscele di reazione sono state allestite in un volume finale di 50 μ l, utilizzando 100 ng di DNA genomico in presenza delle seguenti concentrazioni di reagenti:

Tampone di reazione (Fermentas)	1X
MgCl ₂ (Fermentas)	1,5 mM
dNTPs (Fermentas)	0,2 mM
<i>primer forward</i>	1 μ M
<i>primer reverse</i>	1 μ M
<i>Taq</i> polimerasi (Fermentas)	1,5 U

Le reazioni sono state condotte con il seguente programma termico:

94°C	3 min	
94°C	30 sec	} 35 cicli
52°C	30 sec	
72°C	1 min	
72°C	7 min	

La seconda reazione di amplificazione è stata condotta utilizzando come stampo 1 μ l della prima reazione.

I frammenti amplificati sono stati clonati nel vettore pSCA (*StrataClone PCR kit*, Stratagene) e i cloni ricombinanti sequenziati con i *primer* T3 e T7.

Amplificazione, clonaggio e sequenziamento di regioni genomiche dei geni *GLO-like* di *Habenaria radiata*

Sono state amplificate, clonate e sequenziate le regioni genomiche dei geni *GLO-like* di *H. radiata* *HrGLO1* e *HrGLO2* (Figura 24) che si estendono dal primo al terzo esone. Il DNA di *H. radiata* è stato gentilmente fornito dal Dott. Kanno dell'Università Tohoku di Sendai, Giappone. Le sequenze dei *primer* sono riportate nella Tabella 10.



Figura 23 - Posizione dei *primer* utilizzati per amplificare la regione che si estende dal primo al sesto esone del locus *OrcPI2*.

Nome <i>primer</i>	Sequenza (5'-3')
HSCF	CAAGTGACCTTCTCGAAGCGG
GF	ATAGATATGTAAGCCTTTTAT
GR2	ATAGATATGTAAGCCTTTTAT

Tabella 9 - Sequenza dei *primer* utilizzati per amplificare il locus *OrcPI2*.

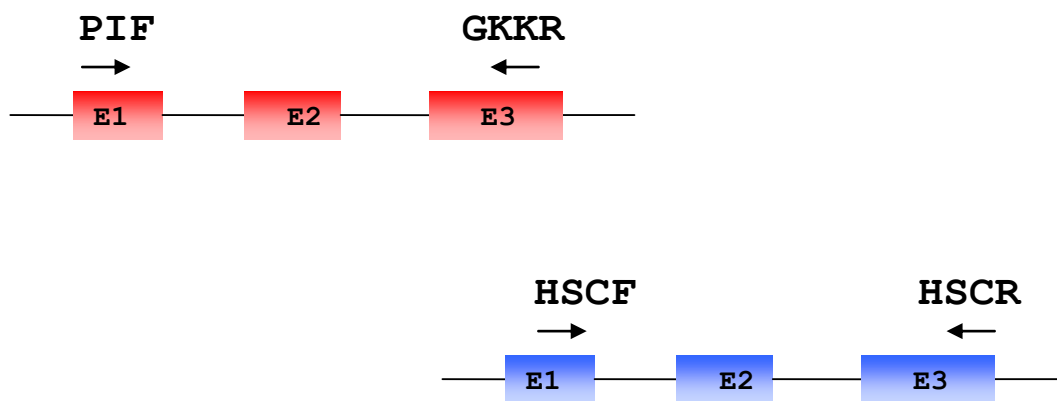


Figura 24 - Posizione dei *primer* utilizzati per amplificare i geni *GLO-like* in *H. radiata*.

Nome	Sequenza (5'-3')
PIF	ATGGGGCGGGGAAATACGGAG
GKRR	TTATCGATCTCCTTCTACCT
HSCF	CAAGTGACCTTCTCGAAGCGG
HSCR	CATGTTATCGTTCTCCTTCTT

Tabella 10 - Sequenza dei *primer* utilizzati per amplificare i geni *GLO-like* in *H.radiata*.

Le miscele di reazione sono state allestite in un volume finale di 50 μ l, utilizzando 100 ng di DNA genomico in presenza delle seguenti concentrazioni di reagenti:

Tampone di reazione (Fermentas)	1X
MgCl ₂ (Fermentas)	1,5 mM
dNTPs (Fermentas)	0,2 mM
<i>primer forward</i>	1 μ M
<i>primer reverse</i>	1 μ M
<i>Taq</i> polimerasi (Fermentas)	1,5 U

Le reazioni sono state condotte con il seguente programma termico:

94°C	3 min	
94°C	30 sec	} 35 cicli
52°C	30 sec	
72°C	1 min	
72°C	7 min	

I frammenti amplificati sono stati clonati nel vettore pSCA (*StrataClone PCR kit*, Stratagene) e i cloni ricombinanti sequenziati con i *primer* T3 e T7.

Analisi delle sequenze

Le sequenze nucleotidiche dei cDNA e le loro traduzioni virtuali sono state confrontate con quelle presenti in banca dati (GenBank) attraverso un'analisi BLAST. Le sequenze aminoacidiche ottenute dalla traduzione virtuale dei cDNA isolati nel presente lavoro sono state allineate con quelle codificate dai geni omologhi presenti in banca dati (Tabella 11) utilizzando il *software* T-Coffee (Notredame et al., 2000) e l'allineamento ottenuto è stato utilizzato come "guida" per allineare le sequenze di cDNA con il *software* BioEdit (Hall, 1999).

Gli allineamenti nucleotidici ottenuti riguardano le tre sub-famiglie geniche *AG-like*, *DEF-like* e *GLO-like* e ciascun allineamento comprende i differenti paraloghi isolati nel presente lavoro e gli omologhi presenti in banca dati.

<u>Geni</u> <u>GLO-like</u>	<u>Specie</u>	<u>Accession</u> <u>number</u>	<u>Geni</u> <u>AG-like</u>	<u>Specie</u>	<u>Accession</u> <u>number</u>	<u>Geni</u> <u>DEF-like</u>	<u>Specie</u>	<u>Accession</u> <u>number</u>
DthyrPI	<i>Dendrobium thyrsiflorum</i>	DQ017701	DenthAG2	<i>Dendrobium thyrsiflorum</i>		MADS1	<i>Cymbidium spp</i>	DQ683575
GogalGLO1	<i>Gongora galeata</i>	FJ804100	PhalAG2	<i>Phalaenopsis equestris</i>	AB232953	GogalDEF1	<i>Gongora galeata</i>	FJ804097
PeMADS6	<i>Phalaenopsis equestris</i>	AY678299	LMADS2	<i>Lilium regale</i>	AY522502	DcOAP3A	<i>Dendrobium crumenatum</i>	DQ119838
EpalPI	<i>Epipactis palustris</i>	DQ005588	DencrAG2	<i>Dendrobium crumenatum</i>	DQ119842	PeMADS2	<i>Phalaenopsis equestris</i>	AY378149
SpodoGLO1	<i>Spiranthes odorata</i>	FJ804114	HoMADS1	<i>Hyacinthus orientalis</i>	AF194335	VaplaDEF1	<i>Vanilla planifolia</i>	FJ804115
HrGLO2	<i>Habenaria radiata</i>	AB232664	ApMADS2	<i>Agapanthus praecox</i>	AB079260	PhlonDEF2	<i>Phragmipedium longifolium</i>	FJ804106
PhlonGLO1	<i>Phragmipedium</i>	FJ804109	AVAG2	<i>Asparagus virgatus</i>	AB175825	SpodoDEF4	<i>Spiranthes odorata</i>	FJ804113
HrGLO1	<i>Habenaria radiata</i>	AB232665	OS	<i>Oryza sativa</i>	AP003379	PhlonDEF1	<i>Phragmipedium longifolium</i>	FJ804105
VaplaGLO1	<i>Vacilla planifolia</i>	FJ804118	ZMM25	<i>Zea mays</i>	AJ430639	SpodoDEF2	<i>Spiranthes odorata</i>	FJ804111
AOGLOA	<i>Asparagus officinalis</i>	AB103465	OSMADS13	<i>Oryza sativa</i>	AF151693	GogalDEF2	<i>Gongora galeata</i>	FJ804099
			ZAG2	<i>Zea mays</i>	L18925	OMADS3	<i>Oncidium "Gower Ramsey"</i>	AY196350
			ZMM1	<i>Zea mays</i>	X81199	PeMADS5	<i>Phalaenopsis equestris</i>	AY378148
			OSMADS3	<i>Oryza sativa</i>	L37528	GogalDEF3	<i>Gongora galeata</i>	FJ804099
			ZAG1	<i>Zea mays</i>	L18924	PeMADS3	<i>Phalaenopsis equestris</i>	AY378150
			AVAG1	<i>Asparagus virgatus</i>	AB125347	DcOAP3B	<i>Dendrobium crumenatum</i>	DQ119839
			HAG1	<i>Hyacinthus orientalis</i>	AF099937	HrDEF	<i>Habenaria radiata</i>	AB232663
			DencrAG1	<i>Dendrobium crumenatum</i>	DQ119841	SpodoDEF1	<i>Spiranthes odorata</i>	FJ804110
			DenthAG1	<i>Dendrobium thyrsiflorum</i>		PhlonDEF3	<i>Phragmipedium longifolium</i>	FJ804107
			PhalAG1	<i>Phalaenopsis equestris</i>	AB232952	VaplaDEF3	<i>Vanilla planifolia</i>	FJ804117
						VaplaDEF2	<i>Vanilla planifolia</i>	FJ804116
						SpodoDEF3	<i>Spiranthes odorata</i>	FJ804112
						PeMADS4	<i>Phalaenopsis equestris</i>	AY378147
						AODEF	<i>Asparagus officinalis</i>	AB094964

Tabella 11 - Elenco dei geni *GLO-like*, *AG-like* e *DEF-like* presenti in GenBank e sottoposti ad analisi evolutiva.

Il software MEGA 4.0.1 (Tamura et al., 2007) è stato utilizzato per condurre un'analisi di distanza Neighbor-Joining (NJ) sui tre differenti allineamenti delle sequenze di cDNA (*AG-like*, *DEF-like* e *GLO-like*), considerando separatamente le intere sequenze e le regioni codificanti i domini MIK (escludendo quindi il dominio C-terminale). La validità statistica degli alberi ottenuti è stata valutata attraverso un'analisi *Bootstrap* condotta con 1000 repliche. Gli allineamenti delle regioni genomiche dei geni *OrcPI* e *OrcPI2* (comprendenti regioni codificanti e non codificanti) delle specie di orchidee analizzate sono stati utilizzati per condurre lo stesso tipo di analisi di distanza, con lo stesso software.

Il software DnaSP 5.0 (Rozas et al., 2003) è stato utilizzato per valutare la variabilità nucleotidica intra e interspecifica delle sequenze del gene *OrcPI* nelle popolazioni di *Anacamptis laxiflora* e di *A. morio* calcolando il polimorfismo nucleotidico (π), la divergenza nucleotidica tra le specie e il valore D di Tajima (Tajima's D) (Tajima, 1989) nei siti codificanti e non codificanti. Il test di Tajima, utilizzando l'ipotesi della neutralità come "ipotesi zero", permette di confrontare il numero di siti segreganti e la diversità nucleotidica osservata. Se tutti gli alleli sono selettivamente neutrali, il prodotto $4N_e\mu$ (definito θ e indicante il prodotto tra il tasso di mutazione per sito, μ , e la dimensione effettiva della popolazione, N_e) può essere espresso in riferimento a due parametri:

- π che esprime la proporzione di differenze nucleotidiche tra coppie di sequenze
- k che esprime il tasso di sostituzione allelica.

Se la variazione nucleotidica è neutrale e la popolazione è in equilibrio, θ_k e θ_π hanno statisticamente lo stesso valore per cui il Tajima's D = $\theta_\pi - \theta_k$ sarà uguale a zero ($D = 0$) (Tajima, 1989).

Se la popolazione ha subito un recente collo di bottiglia, π non ne sarà influenzato, ma k potrebbe essere notevolmente ridotto, quindi D dovrebbe essere maggiore di 0. Se la popolazione ha subito una espansione recente, è poco probabile che le nuove mutazioni vengano perse, ma sarà necessario molto tempo affinché contribuiscano all'eterozigosità θ_k , quindi D sarà negativo.

Il test di Hudson, Kreitman e Aguade (HKA) (Hudson et al., 1987) è stato condotto su differenti combinazioni di esoni e introni, utilizzando i siti sinonimi e i siti silenti (non codificanti). Il test HKA si basa sulla correlazione positiva tra tasso di mutazione (π/K), divergenza tra specie (K) e polimorfismo intraspecifico (π). L'analisi necessita di dati riguardanti almeno due regioni genetiche differenti in due specie differenti e permette, tramite confronto dei tassi di mutazione, di valutare deviazioni dal modello neutrale.

L'allineamento delle sequenze nucleotidiche codificanti dei geni *OrcPI* e *OrcPI2* delle specie di orchidee analizzate è stato utilizzato per calcolare il tasso di sostituzioni sinonime (K_s) e non sinonime (K_a) utilizzando il modello di distanza Jukes-Cantor (Jukes e Cantor, 1969) con il metodo di Nei-Gojobori (Nei e Gojobori 1986) implementato nel software MEGA 4.0.1 (Kumar et al. 2004). E' stato calcolato il rapporto medio K_a/K_s (ω) per verificare se le sequenze di *OrcPI* e *OrcPI2* siano soggette a selezione purificante ($\omega < 1$) positiva ($\omega > 1$) o neutrale ($\omega = 1$).

Partendo dalla considerazione che la selezione positiva può essere localizzata in codoni specifici circondati da regioni sottoposte a forte selezione purificante e non essere quindi facilmente rilevabile dal semplice confronto tra i tassi di sostituzione nucleotidica, è stata condotta l'analisi CODEML utilizzando il software PAML 4.1 (Yang, 1997) sull'allineamento delle regioni codificanti dei geni *OrcPI* e *OrcPI2* delle specie di orchidee analizzate per identificare eventuali variazioni dei tassi evolutivi. In particolare, sono stati confrontati modelli alternativi di sostituzione di codoni che escludono (M1a e M7) e consentono (M2a e M8) la selezione positiva (Yang et al. 2000; Zhang et al. 2005; Yang et al. 2005). Un ulteriore confronto è stato condotto per rilevare eventuali segnali di selezione positiva all'interno del clade comprendente le sequenze del gene *OrcPI* e quello comprendente le sequenze del gene *OrcPI2* (Modello A versus Modello A_nullo).

Le sequenze codificanti complete dei cDNA di *OrcPI/HrGLO1* e *OrcPI2/HrGLO2* di *O. italica* e *H. radiata* sono state utilizzate per condurre il Relative Rate test utilizzando il software RRTree (Robinson-Rechavi e Huchon 2000). L'ipotesi nulla di questo test

assume che i due gruppi di sequenze confrontate (*OrcPI/HrGLO1* e *OrcPI2/HrGLO2*) non accumulino sostituzioni a tassi significativamente differenti. Il test è stato condotto sulle sostituzioni sinonime e su quelle non sinonime.

La sequenza nucleotidica del promotore del gene *OrcPI* è stata confrontata con le sequenze del promotore di alcuni geni omologhi presenti in banca dati per identificare gli elementi *cis*-regolativi conservati utilizzando l'approccio noto come "*phylogenetic footprinting*" (De Bodt et al., 2006). Questo approccio consiste nell'analisi comparativa di sequenze di DNA non codificante per identificare possibili regioni regolative conservate durante l'evoluzione a causa di costrizioni funzionali (Gumucio et al., 1992; Boffelli et al., 2003; Hong et al., 2003). Il razionale di questo tipo di analisi si basa sull'assunzione che le sequenze conservate presenti nelle regioni regolative di geni ortologhi abbiano una rilevanza funzionale. Queste regioni conservate di DNA non codificante possono essere molto corte (6-12 bp) e circondate da ampie regioni non conservate, soprattutto quando si confrontano specie filogeneticamente molto distanti (Inada et al. 2003).

Le sequenze analizzate già presenti in banca dati sono le seguenti:

- la regione non codificante all'estremità 5' (1352 bp a monte del sito di inizio traduzione) dei due geni *GLO-like*, *OSMADS2* (AP008211) e *OSMADS4* (AP008207), di *Oryza sativa* (Chung et al., 1995), l'unica monocotiledone il cui genoma è stato completamente sequenziato. Queste sequenze sono state chiamate *ORY1* e *ORY5*, facendo riferimento alla loro posizione sul cromosoma 1 e 5, rispettivamente;
- la regione non codificante all'estremità 5' (chiamata *LIL*, di 984 bp) del gene *GLO-like LRGLO* (AB158292) di *Lilium regale*, una specie monocotiledone filogeneticamente più vicina alle orchidee rispetto a *O. sativa*;
- una regione (1352 bp) del promotore del gene *PISTILLATA* (*PI*) di *Arabidopsis thaliana* (AF198055), l'unico promotore di *PI* caratterizzato funzionalmente (Chen et al., 2000; Honma e Goto, 2000).

Gli elementi regolativi principali del promotore di *PI* di *A. thaliana* risiedono all'interno delle 1500 bp a monte del sito di inizio della traduzione. Questa sequenza è sufficiente a guidare l'espressione di *PI* nelle piante selvatiche di *Arabidopsis* e a complementare l'allele mutante nullo *pi-2* (Honma e Goto, 2000). Inoltre, la regione di 481 bp a monte del codone di inizio della traduzione del gene *PI* è sufficiente al conferimento della regolazione spaziale e temporale dell'espressione di *PI* (Chen et al., 2000).

Le sequenze nucleotidiche delle regioni del promotore del locus *OrcPI* di *O. italica* e dei loci *GLO-like* di *O. sativa* e *L. regale* sono state analizzate usando il programma *FootPrinter3* (Blanchette e Tompa, 2003; Fang e Blanchette, 2006) che consente di identificare, in un insieme di sequenze genomiche ortologhe o paraloghe, probabili elementi *cis*-regolativi conservati. L'analisi è stata condotta più volte, selezionando la lunghezza dei motivi conservati da 9 a 12 bp e con un numero di mutazioni variabile da 0 a 3. Successivamente è stata condotta l'analisi aggiungendo la sequenza di 1352 bp del promotore del gene *PI* di *A. thaliana*.

Tutte le sequenze dei promotori *GLO-like* precedentemente descritte sono state confrontate con quelle depositate nelle banche dati *Place* (Higo et al., 1999), *PlantCare* (Lescot et al., 2002) e *PlantPAN* (Chang et al., 2008), specifiche per siti di legame dei fattori di trascrizione vegetali. I motivi identificati di lunghezza superiore a 5 bp sono stati confrontati con quelli individuati dal programma *FootPrinter3*.

Risultati e Discussione

Identificazione di geni MADS-box in *Orchis italica*

L'approccio utilizzato per identificare geni MADS in orchidee mediterranee è stato facilitato dalla natura molto conservata della sequenza aminoacidica e nucleotidica del dominio MADS-box. Tale caratteristica ha permesso di allestire, tramite RT-PCR, una mini-libreria di cDNA di geni MADS-box di *O. italica*, la cui analisi ha consentito di isolare 8 differenti loci: due *AG-like* (*OrcAG1* e *OrcAG2*), quattro *DEF-like* (*OrcDEF1*, *OrcDEF2*, *OrcDEF3* e *OrcDEF4*), due *GLO-like* (*OrcPI* e *OrcPI2*).

L'analisi evolutiva si è focalizzata sul locus *GLO-like*, con particolare riguardo al gene *OrcPI*. La scelta di incentrare lo studio su *OrcPI* è dovuta alla peculiare presenza di due copie funzionali (paraloghi) di questo locus solo in alcune specie di orchidee appartenenti alle *Orchidoideae*. Studi recenti (Mondragon-Palomino et al., 2009) hanno riportato il tentativo di identificare la seconda copia del locus *GLO-like* in specie di orchidee appartenenti ad altre sottofamiglie (*Epidendroideae*, *Cypripedioideae*, *Vanilloideae*). Analoghi tentativi sono stati condotti anche durante il mio lavoro di tesi (dati non riportati). In entrambi i casi, gli esiti della ricerca sono stati negativi. La presenza di una seconda copia di questo locus nelle orchidee è descritta solo per *Habenaria radiata* (Kim et al., 2007), specie appartenente alla sottofamiglia *Orchidoideae*, tribù *Orchidinae*.

Analisi dei geni *AG-like* in *O. italica*

Le funzioni svolte dai geni *AG-like* nelle dicotiledoni e nelle monocotiledoni, pur essendo ampiamente conservate, presentano alcune differenze la cui origine è probabilmente riconducibile ad eventi duplicativi. Infatti, lo scenario evolutivo dei geni *AG-like* è stato contrassegnato da duplicazioni seguite da diversificazione funzionale tramite un meccanismo di sub-funzionalizzazione, che sembra aver determinato una restrizione del profilo di espressione dei geni delle classi C e D,

facilitando in questo modo le modificazioni evolutive dei tessuti riproduttivi (Kramer et al., 2004).

Nelle *Orchidaceae*, i geni *AG-like* di classe C e D mostrano un profilo di espressione peculiare: i loro trascritti non sono presenti solo negli organi del terzo e del quarto verticillo (profilo tipico dei geni di classe C e D nelle dicotiledoni), ma si osservano anche nel primo e nel secondo verticillo. E' stato quindi ipotizzato che la modificazione dell'espressione dei loci di classe C e D possa aver contribuito alla formazione degli organi florali caratteristici delle orchidee, come il labello e la colonna, e che i geni di classe D, mostrando un'espressione non confinata solo agli ovuli e un motivo caratteristico all'interno della regione C-terminale (motivo MD) rispetto ai geni di classe C, possano svolgere nuove funzioni all'interno delle orchidee (Skipper et al., 2006; Song et al., 2006; Xu et al., 2006).

L'analisi della mini-libreria di cDNA di geni MADS-box di *O. italica* ha evidenziato la presenza di due differenti cDNA *AG-like* di lunghezza 840 bp (*OrcAG1*) e 740 bp (*OrcAG2*). I due cDNA non rappresentano l'intera lunghezza della regione codificante, mancando di una corta regione all'estremità 5' comprendente il 5'UTR e 40 bp del primo esone. Entrambe le sequenze comprendono la regione 3'UTR. L'analisi BLAST ha rivelato che *OrcAG1* e *OrcAG2* hanno un'elevata similarità nucleotidica, rispettivamente con *DencrAG1* e *DencrAG2* dell'orchidea *Dendrobium crumenatum* (Xu et al., 2006). La traduzione virtuale delle due sequenze nucleotidiche (211 aminoacidi per *OrcAG1*, 218 aminoacidi per *OrcAG2*) ha evidenziato una identità aminoacidica dell'85% tra le due sequenze.

L'allineamento con altre proteine *AG-like* ha evidenziato che *OrcAG1* e *OrcAG2* hanno la tipica struttura modulare MIKC e presentano nella regione C-terminale i due motivi caratteristici della famiglia *AG-like*: i motivi AG I e II (Kramer et al., 2004); inoltre, nella regione C-terminale di *OrcAG2* è presente il motivo MD (Figura 25). Tali motivi sono stati identificati anche nei geni *AG-like* di *Phalaenopsis* (*PhlAG1*, *PhlAG2*), di *D. crumenatum* (*DencrAG1*, *DencrAG2*) e *D. thyrsiflorum* (*DthyrAG1*, *DthyrAG2*) (Figura 25) (Yu et al., 2004; Song et al., 2006; Xu et al., 2006; Skipper et al., 2006). I motivi individuati sono considerati

OrcAG1	NKIADNERSQQQQ-HMNMPSTSTDYEMMP	FDSRNFLQVNLMD	PNSHY	S-LQQQTALQLG
DencrAG1	NKIADNERTQQQH-HINMVPSTSTEYEVMP	FDSRNFLQVNLMD	PSHHY	S-LQQQTALQVG
DenthAG1	NKIADNERAQQHQ-HMNILPSTSAEYEVMP	FDSRSFLQVNLDD	PNDHY	A-HQQQTALQLG
PhalAG1	NKISDNERAQQQHMHMSILPSTSTEYEVMP	FDSRSFLHVNLM	PNDRY	S-HQQQTALQLG
OrcAG2	TKIAENERARAN-----NEESATGFNDLSS	FDSRNYYHVNMF	AAAAH	Y-NQNQTSLHLG
DencrAG2	AKITENERQ--TN-----IDTTASALDTLST	FDSRNYYPVNMLE	AAAHY	H-NQDQTALHLG
DenthAG2	AKINDNERAEQAN-----IVQAGADFDTLPN	FDSRNYYQVNILE	TAAHY	SHHQDQTALHLG
PhalAG2	AKIADNERAQQAN-----IVQAGVDFESIPS	FDSRNYYHINMLE	SASHY	SHHQDQTALHLG

OrcAG1	-----		
DencrAG1	-----		
DenthAG1	-----		
PhalAG1	-----		
OrcAG2	YDTKVE---		
DencrAG2	YDN-----		
DenthAG2	YETKADHSA		
PhalAG2	YETKADQSA		

		AG-motif 1	AG-motif 2
--	--	------------	------------

	MD-motif		
--	----------	--	--

Figura 25 - Allineamento della regione C-terminale delle sequenze aminoacidiche delle proteine AG-like di *Orchidaceae*.

fondamentali per la funzionalità delle proteine AG-like e tale idea è supportata da diverse evidenze sperimentali (Kramer et al., 2004); in particolare, è stato ipotizzato che il motivo MD, una sequenza di 9 aminoacidi a valle del motivo AG II, tipico delle proteine AG-like di classe D delle monocotiledoni, possa svolgere in questo gruppo di piante un ruolo critico per la funzionalità delle proteine influenzando, ad esempio, l'interazione proteina-proteina, la specificità di legame al DNA, la localizzazione sub-cellulare e di conseguenza contribuire all'evoluzione fenotipica delle monocotiledoni (Yu et al., 2004; Song et al., 2006).

L'analisi di distanza è stata condotta sull'allineamento delle sequenze dei cDNA dei geni *OrcAG1* e *OrcAG2* e delle sequenze dei cDNA AG-like di diverse orchidee presenti in banca dati. L'analisi è stata condotta separatamente su due allineamenti differenti, uno comprendente l'intera sequenza codificante e l'altro comprendente solo la regione codificante per i domini MIK (escludendo la regione codificante il dominio C-terminale). Tale scelta è dovuta alla elevata variabilità di sequenza della regione codificante il dominio C-terminale che rende difficoltoso e ambiguo l'allineamento nucleotidico. L'albero ottenuto dall'analisi dell'allineamento dell'intera regione codificante, infatti, presenta numerosi rami non risolti (collassati) e i cladi risolti presentano percentuali di *Bootstrap* non elevate. L'analisi condotta sull'allineamento delle regioni codificanti i domini MIK ha prodotto un albero (Figura 26) in cui si evidenziano due cladi ben distinti, corrispondenti alle classi C e D. *OrcAG1* si posiziona all'interno della classe C, mentre *OrcAG2* nella classe D. Ulteriori studi comparativi, sia funzionali sia di evoluzione molecolare, sono necessari per chiarire il ruolo dei geni AG-like nella diversificazione del fiore delle *Orchidaceae*.

Analisi dei geni DEF-like in *O. italica*

L'importanza evolutiva della regione C-terminale delle proteine MADS-box è supportata da numerosi studi che suggeriscono un ruolo fondamentale delle mutazioni *frameshift* del dominio C-terminale nella divergenza strutturale e funzionale delle proteine MADS-box influenzando, di conseguenza, lo sviluppo fiorale (Kramer et al.,

- GENI DI CLASSE D
- GENI DI CLASSE C

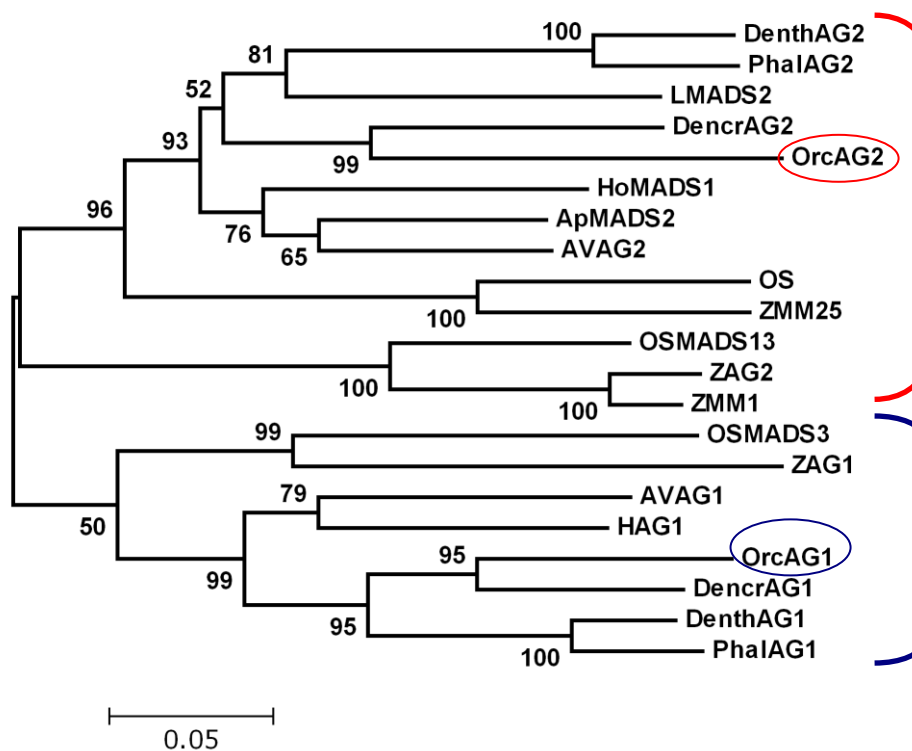


Figura 26 - Albero di distanza N-J ottenuto dall'analisi dell'allineamento delle regioni codificanti i domini MIK dei geni AG-like. I numeri rappresentano i valori di *Bootstrap* espressi in percentuale.

2006). Ad esempio, nella famiglia *AP3/DEF* è possibile distinguere le due sottofamiglie *euAP3* e *TM6*, i cui membri presentano motivi differenti all'interno della regione C-terminale: il motivo *euAP3* e il motivo *paleoAP3*, rispettivamente. Il motivo *euAP3* (presente solo nelle proteine *DEF-like* delle eudicotiledoni superiori) ha avuto origine da una mutazione *frameshift* (inserzione di otto coppie di basi) all'interno del motivo *paleoAP3* (Kramer et al., 2006). Inoltre, l'analisi di mutanti di *A. thaliana* ha dimostrato che i due motivi conferiscono funzioni differenti (Lamb e Irish, 2003). E' stato ipotizzato quindi che la diversificazione funzionale delle proteine *DEF-like*, guidata da mutazioni *frameshift* all'interno della regione C-terminale, abbia contribuito alla diversificazione fiorale delle angiosperme (Kramer et al., 2006).

Un recente studio evolutivo condotto sui geni *DEF-like* nelle *Orchidaceae* ha evidenziato che i membri dei quattro cladi di questa sottofamiglia genica (cladi 1-4) hanno funzioni diverse. La diversificazione funzionale viene fatta risalire ad eventi duplicativi seguiti da fenomeni di selezione positiva localizzata in siti specifici del dominio K e del dominio C-terminale (Mondragon-Palomino et al., 2009). Il dominio C-terminale è caratterizzato da una marcata diversificazione clade-specifica (Mondragon-Palomino et al., 2009). L'identità nucleotidica relativamente elevata (68-92%) della regione MIK dei geni dei cladi 1, 2, 3 e 4 tende a diminuire nella regione C-terminale (46-90%) e la divergenza nucleotidica è particolarmente pronunciata nel clade 2 (28-57%) a causa di numerose sostituzioni, di codoni di stop prematuri e di delezioni specie-specifiche (Mondragon-Palomino et al., 2009). E' stata ipotizzata quindi l'esistenza di una connessione tra duplicazione dei geni *DEF-like*, mutazioni della regione che codifica il dominio C-terminale, diversificazione funzionale dei geni *DEF-like* ed evoluzione di nuove caratteristiche fiorali nelle orchidee (Mondragon-Palomino et al., 2009).

L'analisi della mini-libreria di cDNA di geni MADS-box di *O. italica* ha evidenziato la presenza di quattro differenti cDNA *DEF-like* di lunghezza 801 bp (*OrcDEF1*), 874 bp (*OrcDEF2*), 729 bp (*OrcDEF3*) e 813 bp (*OrcDEF4*). I quattro cDNA rappresentano

l'intera lunghezza della regione codificante e comprendono la regione 3'UTR, ma non la 5'UTR.

La traduzione virtuale delle sequenze in esame produce proteine di diversa lunghezza: 250 aa (OrcDEF1), 216 aa (OrcDEF2), 240 aa (OrcDEF3) e 241 aa (OrcDEF4), con la tipica struttura modulare MIKC.

A differenza del dominio MADS, che risulta essere molto conservato, nella regione IKC delle proteine OrcDEF si osserva una grande variabilità; in particolare, nella regione C-terminale le sequenze aminoacidiche OrcDEF1, OrcDEF3 e OrcDEF4 mostrano una similarità del 60%, mentre la sequenza di OrcDEF2 è molto più divergente. All'interno della regione K, i geni *DEF-like* presentano un motivo conservato, la cui sequenza è (H/Q)YEXM (Kramer et al., 1998); questo motivo è stato individuato anche nei quattro paraloghi OrcDEF in posizione conservata, ma nelle copie OrcDEF1 e OrcDEF2 è presente un residuo F nella seconda posizione del motivo (QFERM), mentre in OrcDEF3 e OrcDEF4 il motivo è QYEKM (Figura 27). Nella regione C-terminale delle proteine OrcDEF, è stato inoltre individuato il motivo *PI-derived* (Kramer et al., 1998; Moon et al., 1999) e il motivo *paleoAP3*; quest'ultimo non è presente all'interno della regione C-terminale di OrcDEF2 (Figura 27) (Kramer et al., 1998).

Per i geni *DEF-like*, così come per i geni *AG-like*, l'analisi di distanza dell'allineamento delle regioni codificanti i domini MIK (escludendo la regione codificante il dominio C-terminale) ha prodotto un albero con rami meglio risolti e statisticamente supportati rispetto all'albero ottenuto dall'analisi di distanza dell'intera regione codificante. Anche in questo caso, la migliore risoluzione è dovuta all'estrema variabilità della regione codificante il dominio C-terminale nei geni *DEF-like*. L'albero ottenuto conferma l'esistenza di quattro cladi *DEF-like* specifici delle orchidee: i geni *PeMADS2-like* (clade 1, comprendente *OrcDEF1*), i geni *OMADS3-like* (clade 2, comprendente *OrcDEF2*), i geni *PeMADS3-like* (clade 3, comprendente *OrcDEF3*) e i geni *PeMADS4-like* (clade 4, comprendente *OrcDEF4*). I cladi 1 e 2 condividono un progenitore comune (sono *sister group*), così come i cladi 3 e 4 (Figura 28). Sarà interessante studiare i profili di espressione e l'evoluzione molecolare dei loci *OrcDEF*, in

A)

PhlonDEF2	STDIKGIFERYQIVTGIELWNAQYERMORTLKHLKDINQSLRKEIMQRTGEGLEGMDIEE
SpodoDEF4	-----QIVTGMELWNAQYERMORTLKHLKDINQSLRKEIMQRTGEGLEGMDIEE
MADS1_Cymb	STDIKGIYERYQIVTGMDLWNAQYERMONTLNHLKEINQNLKEIRQRNGEELEGLDIKE
GogalDEF1	STDFKGLYERYQIVTGVDLWNAQYERMONTLRHLKEINQNLKEIRQRKGEELEAMDIEE
PeMADS2	STDIKGIYERYQVVTGMDLWNAQYERMONTLKHLKEINQNLKEIRRRKGEELEGMIDIKQ
DcOAP3A	STDIKGVYERYQVVTGIDLWNAQYERMONTLKHLKEINQNLKEIRQRKGEELEGMIEKE
OrcDEF1	STDVKATYERYQAITGIDLWNAQFERMONTLKHLQEIINQNLREIRQRNGKELEGLDIME
VaplaDEF1	STDIKAIFEKYQLVTGDLWNAQYERMONTLKLLKEMNQKLREISQRKGEDLDGMDIKE
SpodoDEF3	STESKKVFDYRQQVSGINLWSAQYEKLONTLNHLKEINHNLKRELQWTGEELEGLDIKE
OrcDEF4	SSETKKIFDRYQVSGINLWSSQYEKMONTLNHLKEINHNLRRDVRQRTGGDLEGLDIKE
PhlonDEF4	STDAKKVFDYRQQVSGINLWSSQYEKMONTLNHLKEINHSLRREIRQRMGDDLEALDIKE
VaplaDEF2	STDTKKVFDYRQQVSSISLWNAQYEKMONTLNHLKDINHNLRRREIRQRMGEDLDGLDIKE
PeMADS4	STETKKVFFERYQVSGINLWSSQYEKMLNTLNHSKEINRNLRRREVRQRMGEDLEGLDIKE
HrDEF	STDTKSVYDRYQVSGVNLWSAQYEKMONTLNHLKEINDNLRRREIRQRMGEDLDGLDIKE
OrcDEF3	STDTKSVYDRYQVSGVNLWSAQYEKMONTLNHLKEINHSLRREIRQRMGEDLDGLDIKE
GogalDEF3	TTDTKSIYDRYQVSGINLWSAQYEKMONTLNHLKEINQNLRRREIRQRMGEDLDGLEIKE
PeMADS3	TTDTKSVYDRYQVSGINLWSEQYEKMONTLNHLKEINHNLRRREIRQRMGEDLEGLEIKE
DcOAP3B	STDTKSIYDRYQQLSGINLWSAQYEKMONTLNHLKEINHNLRRREIRQRMGEDLDGLEIKE
SpodoDEF1	STDTKSIYDRYQVSGINLWSSQYEKMONTLNHLKEINHNLRRREIRQRTGEDLEGLDVKE
PhlonDEF3	STDAKNMYDRYQATGIDLWSSQYERMONTLSHLKDVNHSRREIRRRMGEDLDGMDIKE
VaplaDEF3	STDTKSVYDRYQHVSIGIDLWSAQYEKMONTLNHLKEINHNLRRREIRQRMGEDLDGLDIKE
GogalDEF2	STEIKDVFFERYQVGTGIDIWDAQYQRMQDTLKNLKEINHNLQKEIRQRKGENLEGLEIKE
OMADS3_Onc	STEIKDAFQRYQVGTGFDIWDAAQYQRMQSTLMNLREVNHKLQMEIRQRKGENLEGLDVKE
PeMADS5	STDVKDIVERYQNVGTGIDIWDAQYQRMONTLRNLREINRNLQKEIRQRKGENLEGLGVKE
OrcDEF2	STDIKGILEQYQVGTGIDIWNAQFERMONILKNLKEINRNLRRREIRQRNGENLEELDILE
SpodoDEF2	SPDINEILHRYQKVTGIDIWHAQEYERMONTLKDLNEINQKLRRSEIRQRIGENLDELDIKE
PhlonDEF1	STNVTQIFRSYHMTGIDIWNSEYERMOKTLKHLKEVNSQLRMEIRQRSGEGLFLNLEE

B)

PhlonDEF2	V-AEEHNRVCDSTIPMVN-VV-PPMYVFRVVQPIQSNLMGLSYN-SHDLSLA
SpodoDEF4	-----
MADS1_Cymb	S-AENHSRVYQNSIPMAT-EC-PQMFSTRVVQPTQPNLLGLGYE-SHDLSLA
GogalDEF1	G-AENHSRMYENSIPMAT-EC-PHMFSTRVAQPIQPSLIGLGYE-SHDLSLA
PeMADS2	N-VENQSRIYENSIPMVN-EC-PQMFSTRVVHPNQPNLLGLGYE-SHDLSLA
DcOAP3A	S-AENHSRVYENSIPMVNNDC-PQMFSTRVVQPIQPNFLIGLGYE-SHDLSLA
OrcDEF1	D-AEDHSRMYENSIPQLLN-NG-PQMLSTRVVQPSQPNLLIGISYD-SHDLSLA
VaplaDEF1	V-AQDPSSVYNGSMIPMVN-GG-PPMFAFRVVQPSQPNLRGVGYE-PHDLNLA
SpodoDEF3	A-D-DVSTNYESAVALAN-GP-SQLYSFR-IQPSHPNLHEMGYG-SHDLRLA
OrcDEF4	V-N-DD-SSYEGALALAH-DGSHQMYTFR-VQPSQPNLHEMGYA-SHDLRLA
PhlonDEF4	I-D-DDPSNYEGTIALAN-GG-SHMYAFR-VQPSQPNLHGMGFG-SHDLRLA
VaplaDEF2	V-E-DDPTNYEGALGLGN-GG-PQMYAFR-VQPNQPNLHAMGYSSSHDLRLA
PeMADS4	H---EDSSNYEGVIALAN-DG-SHMYAFR-VQPNQQNLQGTGYS-SHDLRLA
HrDEF	V-DEDPAGSYDGALALGN-GT-PYLY-FR-TQPTHPNLHGMGYG-SHDLRLA
OrcDEF3	V-DEDPAGSYDGALALGN-GT-PYLYSFR-TQPTQPNLQGMGYG-SHDLRLA
GogalDEF3	V-D-DDPNNYDGALALGN-GS-SYLYSYR-TQPSQPNLQGMGYG-PHDLRLA
PeMADS3	V-D-DDPNNYDGALALGN-GA-SYLYSFR-TQPSQPNLQGVGYV-PHDLRLA
DcOAP3B	V-D-DDPSNYDGALALGN-GA-SYLYSYR-TQPSQPNLQGMGYG-PHDLRLA
SpodoDEF1	V-D-DDPNNYDGALALGN-GG-SYLYSFR-PQPSQPNLQGMGYG-SHDLRLA
PhlonDEF3	V-D-DDQCNDGRLALVN-EA-SYIYSLR-TQPSHTNLIGLGYG-SHDLRLA
VaplaDEF3	V-E-PNPSGYDGRILALTN-GV-SHVYAFR-ADPTQPSLHGVGYD-SHDLRLA
GogalDEF2	V-AEDLIGVYDSAILMAN-QQ-R-----T-VSQICRM
OMADS3_Onc	I-AEDLSGVYNSAISMAN-----QRLAHCL---
PeMADS5	L-VEDLSCIYDSSISMAN-RL-H-----RSEPNVQKVVRACHEFGFD
OrcDEF2	G-AEDLSGVYDGSISMVN-LQ-----H---DE
SpodoDEF2	D-ADDLLYNED-----DRIGLVYE-SHDLN-F
PhlonDEF1	ANSEGLSGVYESLCSMVN-GR-H---NSE-VQPSELSLQGITVYG-WHDL-IF

PI-derived motif Paleo-AP3 motif

Figura 27 - Allineamento della regione K (A) e C-terminale (B) delle proteine DEF-like di *Orchidaceae*.

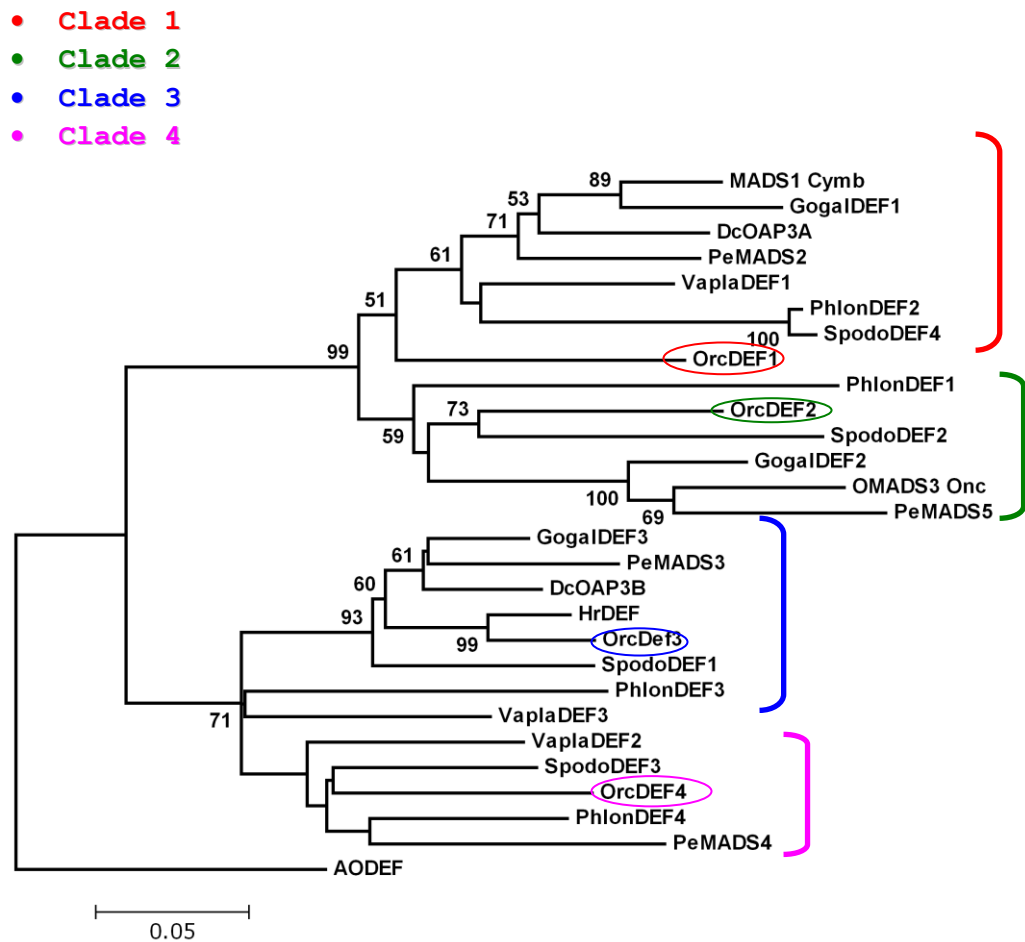


Figura 28 - Albero di distanza N-J ottenuto dall'allineamento delle regioni codificanti i domini MIK dei geni *DEF-like*. I numeri rappresentano i valori di *Bootstrap* espressi in percentuale.

particolare del gene *OrcDEF2*, per comprendere se la diversità nucleotidica e aminoacidica possa essere correlata a diversità morfologiche.

Analisi dei geni *GLO-like* in *O. italica*

Come detto nei paragrafi precedenti, nelle orchidee i geni *GLO-like* sono a singola copia, fatta eccezione per *H. radiata* (*Orchidoideae*, *Orchidinae*) in cui ne sono state isolate due copie (*HrGLO1*, *HrGLO2*) (Kim et al., 2007).

I geni *GLO-like* in *H. radiata* presentano differenze nel profilo di espressione: *HrGLO1* è espresso in tutti gli organi fiorali, *HrGLO2* è espresso nei tepali interni, meno nei tepali esterni e non si osserva espressione nella colonna (Kim et al., 2007). Inoltre, il confronto delle sequenze aminoacidiche di *HrGLO1* e *HrGLO2* ha messo in evidenza una divergenza localizzata all'interno del dominio MADS e del dominio K. E' stata quindi ipotizzata una possibile diversificazione funzionale delle due copie attribuibile ai cambiamenti del dominio MADS e del dominio K, siti che sembrano svolgere un ruolo importante nell'acquisizione di nuove funzioni (Martinez-Castilla et al., 2003).

L'analisi della mini-libreria di cDNA di geni MADS-box ha evidenziato la presenza anche in *O. italica* (*Orchidoideae*, *Orchidinae*) di due geni *GLO-like*, chiamati *OrcPI* e *OrcPI2*. Il cDNA di *OrcPI* è lungo 910 bp e quello di *OrcPI2* 912 bp. Entrambi i cDNA comprendono la sequenza codificante completa e le regioni 5' e 3'UTR.

L'analisi BLAST indica che le sequenze dei cDNA di *OrcPI* e di *OrcPI2* sono molto simili a quelle di *HrGLO1* e *HrGLO2*, rispettivamente, di *H. radiata*. La traduzione virtuale di *OrcPI* e di *OrcPI2* produce in entrambi i casi una sequenza aminoacidica di 210 amminoacidi (Figura 29). Le percentuali di identità tra le sequenze aminoacidiche sono riportate nella Tabella 12. L'allineamento delle sequenze aminoacidiche delle proteine *OrcPI* e *OrcPI2* con altre proteine MADS-box ha messo in evidenza che esse hanno la tipica struttura MIKC.

OrcPI	MGRGNTEIKR	IENSTNRQVT	FSKRRSGIHK	KAREISVLCD	AQVSLVIFSS	LGKLSEYCSF
HrGLO1	MGRGNTEIKR	IENSTNRQVT	FSKRRSGIHK	KAREISVLCD	AQVSLVIFSS	LGKLSEYCSF
OrcPI2	MGRGKIEIKR	IENSTNRQVT	FSKRRNGIMK	KAKEISVLCD	AQVSLVIFSS	LGKMFYCSF
HrGLO2	MGRGKIEIKR	IENSTNRQVT	FSKRRNGIMK	KAKEISVLCD	AQVSLVIFSS	LGKMFYCSF
OrcPI	STTLKMLER	YQQNSGKKLW	DATHENLSAE	IDRIKKENDN	MQIELRHLKG	EDLNSLNPKE
HrGLO1	STTLKMLER	YQQNSGKKLW	DATHENLSAE	IDRIKKENDN	MQIELRHLKG	EDLNSLNPKE
OrcPI2	STSLKMLEK	YQQNSGKKLW	DAKHENLSTE	IDRIKKENDN	MQIEPRHLKG	EDLNSLNPKE
HrGLO2	STSLKMLEK	YQQNSGKKLW	DAKHENLSAE	IDRIKKENDN	MQIELRHLKG	EDLNSLNPKE
OrcPI	LIPIEEGLQN	GLTSVREKQM	DFLKMLRKNE	RMLEENKRL	KYLLQHQQLA	IEGSMRELEI
HrGLO1	LIPIEEALQN	GLTSVREKKK	DFLKMLKKNE	RMLEENKRL	KYLLHHQQMA	IEGSMRELEI
OrcPI2	LIPIEEALQN	GLAGVRDKQM	DFLKMLKKNE	RMLEENKRL	TYLLHHQQLA	MEGSMRELDI
HrGLO2	LIPIEEALQN	GLAGVRDKQS	DFLKMLKKKE	RMLEENKRL	TYLLHHQQLA	MEGSMRELDI
OrcPI	SYHQKDPEYA	--DMPMTFRV	QPFQPNLHGN	N		
HrGLO1	SYHQKDAEYA	--NMPMIFRV	QPFQPNLHGN	N		
OrcPI2	GYHQKDREYA	PPMPMTFRV	QPIQPNLQGN	K		
HrGLO2	G--QKDREYA	PPMPMTFRV	QPIQPNLQGN	K		

PI-motif

Figura 29 - Allineamento delle sequenze aminoacidiche delle proteine GLO-like di *O. italica* e *H. radiata*.

Geni	cDNA tot	Esoni	Introne 1 e 2	Traduzione virtuale
<i>OrcPI</i> vs <i>HrGLO1</i>	79.6%	94.9%	88.5%	95.2%
<i>OrcPI2</i> vs <i>HrGLO2</i>	85.9%	94.3%	85.1%	97.2%
<i>OrcPI</i> vs <i>OrcPI2</i>	68.3%	84.6%	42.1%	86.3%
<i>HrGLO1</i> vs <i>HrGLO2</i>	71.6%	84.6%	58.9%	84.8%

Tabella 12 - Confronto (riportato come percentuale di identità nucleotidica e aminoacidica) tra i geni *GLO-like* di *O. italica* e *H. radiata*.

La regione C-terminale delle proteine *GLO-like* contiene un motivo conservato, noto come motivo PI (Kramer et al., 1998), riscontrato sia in *OrcPI* sia in *OrcPI2* (Figura 29).

La topologia dell'albero di distanza costruito sulla base dell'allineamento dell'intera sequenza codificante dei geni *GLO-like* delle specie di orchidee presenti in banca dati e delle specie analizzate nel presente lavoro evidenzia che i geni *GLO-like* delle orchidee formano due diversi gruppi monofiletici. Un gruppo comprende la maggior parte dei geni *GLO-like* (inclusi i geni *HrGLO2* e *OrcPI2*), con topologia in generale accordo con la filogenesi delle specie analizzate; l'altro gruppo comprende *OrcPI*, *HrGLO1* (*Orchidoideae*, *Orchidinae*) e *PhragGLO1* (*Cypripedioideae*) (Figura 30). Sulla base della topologia di questo albero si potrebbe ipotizzare che le due copie *GLO-like* presenti nelle *Orchidoideae* abbiano avuto origine da una duplicazione avvenuta prima della divergenza tra *Cypripedioideae* e *Orchidoideae*, seguita dalla perdita nella maggior parte delle sottofamiglie di una delle due copie dei geni duplicati durante l'evoluzione delle orchidee. Un'ipotesi alternativa potrebbe essere una duplicazione recente avvenuta all'interno delle *Orchidineae*, anche considerando la posizione statisticamente non ben supportata di *PhlonGLO1* (percentuale di *Bootstrap* 53%) rispetto a *OrcPI* e *HrGLO1* (percentuale di *Bootstrap* 100%) (Figura 30).

L'analisi di distanza dell'allineamento delle sequenze genomiche del gene *OrcPI* delle orchidee analizzate nel presente lavoro ha prodotto un albero con una buona risoluzione dei rami e buoni valori di *Bootstrap* quando alle regioni codificanti venivano aggiunte le sequenze dell'introne 2 (Figura 31). Anche per l'allineamento delle sequenze genomiche del gene *OrcPI2* si ha un risultato simile, nel senso che la risoluzione migliore dei rami si ottiene quando alla regione codificante sono aggiunte le sequenze degli introni (da 1 a 5) (Figura 32).

Per verificare la presenza di tracce di selezione positiva nelle sequenze dei geni *OrcPI* e *OrcPI2* delle varie specie di *Orchidoideae* in esame, l'allineamento delle sequenze codificanti è stato analizzato utilizzando il software PAML. I modelli di sostituzione di codoni che consentono selezione positiva (M2a e

- Geni GLO1
- Geni GLO2

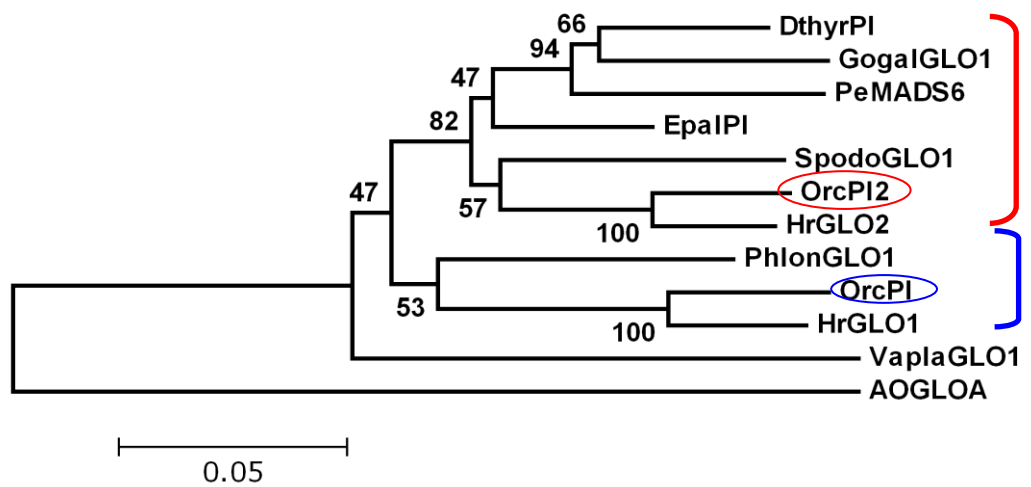


Figura 30 - Albero di distanza N-J ottenuto dall'analisi dell'allineamento delle regioni codificanti dei geni *GLO-like* in *Orchidaceae*.

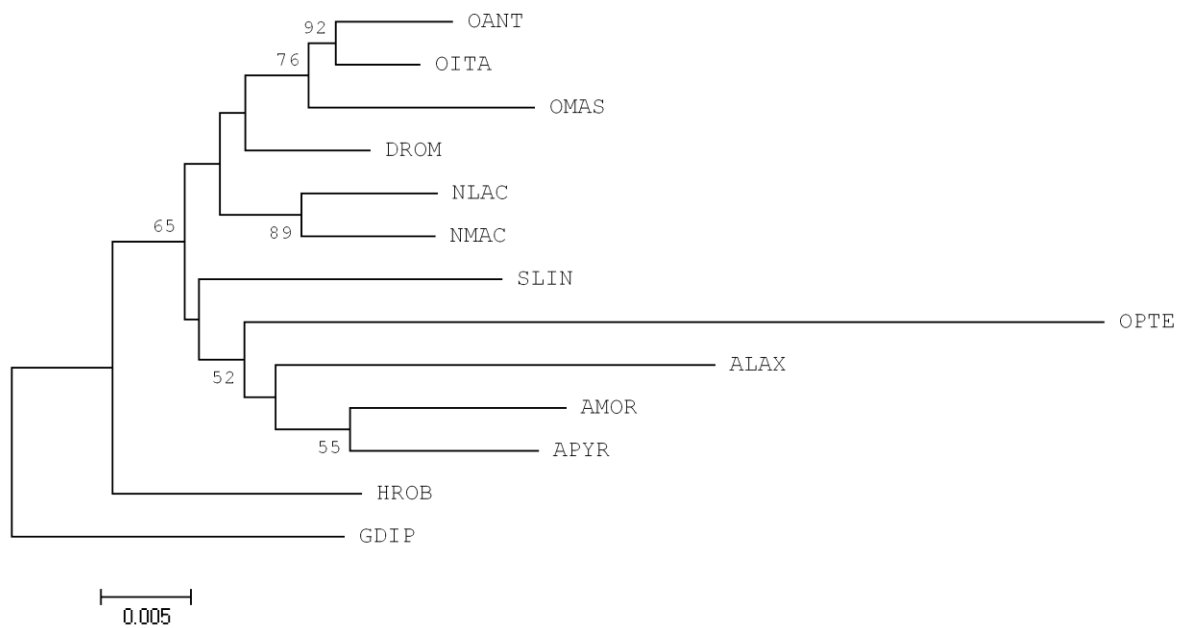


Figura 31 - Albero di distanza N-J ottenuto dall'analisi dell'allineamento delle sequenze codificanti e dell'introne 2 di *OrcPI* nelle *Orchidaceae* analizzate.

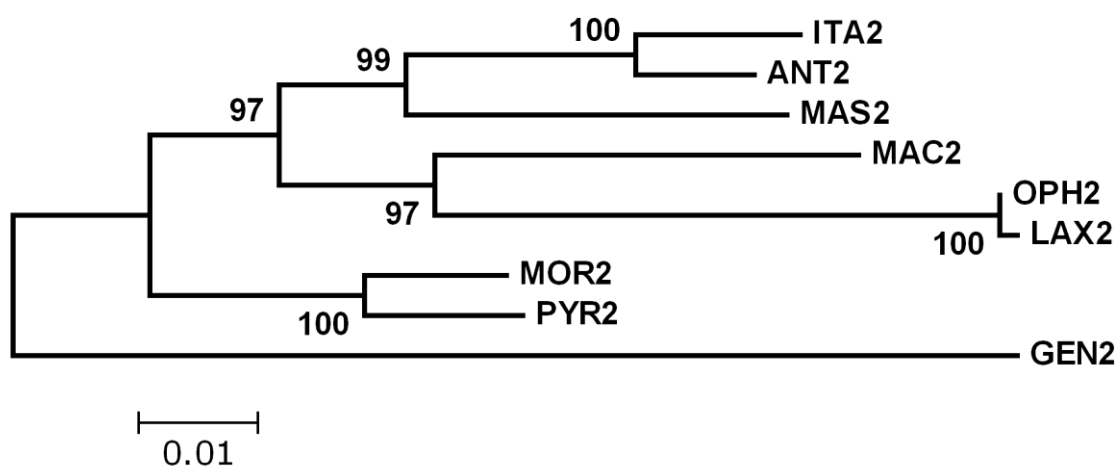


Figura 32 - Albero di distanza ottenuto dall'analisi dell'allineamento della regione genomica che si estende dall'esone 1 all'esone 6 di *OrcPI2* nelle *Orchidaceae* analizzate.

M8) e il modello A di selezione positiva nel clade *OrcPI* rispetto al clade *OrcPI2* non sono in accordo con i dati analizzati e non è stato individuato nessun aminoacido soggetto a selezione positiva.

Il calcolo del tasso di sostituzione sinonima e non sinonima tra i geni *GLO-like* di *O. italica* e *H. radiata* ha evidenziato che sia *OrcPI/HrGLO1* sia *OrcPI2/HrGLO2* presentano valori di K_a (tasso di sostituzioni non sinonime) e K_s (tasso di sostituzioni sinonime) non omogenei (Tabella 13). Il confronto *OrcPI/HrGLO1* ha evidenziato un tasso minore di sostituzioni sinonime e un tasso maggiore di sostituzione non sinonime rispetto al confronto *OrcPI2/HrGLO2*. Sebbene sia evidente che su entrambe le copie dei geni *GLO-like* agisca una diffusa selezione purificante ($K_a/K_s \ll 1$), questo risultato evidenzia che i geni *OrcPI/HrGLO1* si evolvono a un tasso più elevato dei geni *OrcPI2/HrGLO2*. L'analisi della significatività della differenza tra i tassi evolutivi dei due paraloghi in *O. italica* e *H. radiata* (*relative rate test*) ha evidenziato che la differenza tra i tassi di sostituzione sinonima non è significativa, mentre è significativa ($P = 0,03$) la differenza tra i tassi di sostituzione non sinonima. L'insieme delle analisi evolutive condotte ha evidenziato che sui geni *OrcPI/HrGLO1* e *OrcPI2/HrGLO2* agiscono costrizioni selettive di intensità differente e che l'evoluzione delle copie duplicate *OrcPI* e *HrGLO1* è guidata da una riduzione dell'intensità della selezione purificante sui siti non sinonimi piuttosto che dalla selezione positiva.

Il locus *OrcPI*

Il gene *OrcPI* è, tra i due paraloghi del locus *GLO-like* delle *Orchidoideae*, la copia "sovrannumeraria", mentre *OrcPI2* è la copia ortologa dei geni *GLO-like* presenti in tutte le orchidee. L'attenzione si è quindi focalizzata sul gene *OrcPI* al fine di comprendere le dinamiche evolutive che agiscono sull'intero locus.

In prima istanza, è stato isolato il promotore di *OrcPI*, la cui sequenza è stata analizzata per identificare i possibili elementi *cis*-regolativi, utilizzando il "*phylogenetic footprinting*" e il confronto con le sequenze presenti in banche dati specifiche per elementi regolative vegetali; in parallelo, sono stati allestiti

Geni	Tasso di sostituzione sinonima	Tasso di sostituzione nonsinonima	Ka/Ks
<i>OrcPI</i> vs <i>HrGLO1</i>	0.1828 ± 0.0408	0.0204 ± 0.0061	0.1116
<i>OrcPI2</i> vs <i>HrGLO2</i>	0.2784 ± 0.0619	0.0111 ± 0.0056	0.0399
<i>OrcPI</i> vs <i>OrcPI2</i>	0.8381 ± 0.1680	0.0631 ± 0.0126	0.0753
<i>HrGLO1</i> vs <i>HrGLO2</i>	0.6600 ± 0.1132	0.0701 ± 0.0129	0.1062

Tabella 13 – Stima del tasso di sostituzione sinonima e non sinonima dei geni *OrcPI* e *OrcPI2* di *O. italica* e degli ortologi *HrGLO1* e *HrGLO2* di *H. radiata*

saggi di espressione *in vivo* per confermare il ruolo funzionale della regione identificata. Si è poi passati alla caratterizzazione genomica della regione codificante (esoni e introni). Dopo averne determinato le caratteristiche strutturali, è stata condotta un'analisi evolutiva confrontando le sequenze nucleotidiche delle regioni codificanti (esoni) e non codificanti (introni, promotore, 5' e 3'UTR) del gene *OrcPI* in varie specie di orchidee mediterranee. Infine è stato analizzato il profilo di espressione di *OrcPI* in differenti tessuti di *O. italica*.

Analisi del promotore del gene *OrcPI*

L'iter sperimentale per l'identificazione del promotore del gene *OrcPI* utilizzando la TAIL-PCR è stato alquanto laborioso. Sono state apportate diverse modifiche alla procedura originale. Innanzitutto, l'attenzione è stata rivolta ai parametri che influenzano l'efficienza della PCR: concentrazione dei *primer*, volume finale di reazione, programma termico di amplificazione. E' riportato di seguito l'elenco dettagliato degli accorgimenti introdotti:

- aumento della concentrazione dei *primer* aspecifici; è stato osservato infatti che dall'amplificazione si otteneva una grande quantità di prodotti aspecifici con sequenze terminali corrispondenti ai *primer* specifici; tale risultato era causato da una concentrazione inadeguata dei *primer* aspecifici;

- riduzione del volume finale di reazione;
- aumento del tempo di estensione;
- riduzione del numero di cicli e modificazione di alcune condizioni termiche.

Grazie alle variazioni effettuate, è stato possibile isolare un frammento di lunghezza 1324 bp a monte del codone ATG di inizio traduzione del gene *OrcPI*.

Il saggio istochimico della β -glucoronidasi ha dimostrato che l'intera regione isolata all'estremità 5' di *OrcPI* (P1324*OrcPI*) e la sua delezione di 854 bp (Δ P854*OrcPI*) inducono l'espressione del gene *GUS* in petali di *R. hybrida* due giorni dopo l'agro-infiltrazione; gli altri due frammenti deleti, Δ P577*OrcPI* e Δ P356*OrcPI*, inducono espressione del gene *GUS* solo dopo 5 giorni

di incubazione (Figura 33). Tali risultati confermano la presenza di elementi regolativi all'interno della regione isolata all'estremità 5' del gene *OrcPI* in grado di indurre l'espressione del gene reporter *GUS*. La differenza temporale riscontrata nell'induzione dell'espressione del gene *GUS* tra i vari tipi di costrutti potrebbe far arguire che nel promotore del gene *OrcPI* siano presenti segnali regolativi "precoci" e "tardivi", come nel promotore del gene *PI* di *A. thaliana* (Chen et al., 2000). Ulteriori studi sono necessari per confermare tale ipotesi.

La sequenza del promotore di *OrcPI* è stata confrontata con le sequenze omologhe di *O. sativa* (*ORY1* e *ORY5*), di *L. regale* (*LIL*) e di *A. thaliana* (*PI*) per identificare eventuali elementi cis-regolativi conservati (*phylogenetic footprinting*) utilizzando il software FootPrinter3 (De Bodt et al., 2006); inoltre, tutte le sequenze sono state confrontate con quelle presenti nelle banche dati *PlantPan*, *PlantCare* e *PLACE*, specifiche per elementi regolativi vegetali.

Le analisi condotte hanno identificato vari motivi che possono essere suddivisi in due gruppi: elementi tipici dei promotori (TATA-box, CAAT-box, G-box e I-box) ed elementi tipici dei promotori dei geni MADS-box. In particolare, nel promotore di *OrcPI* è stato individuato, in posizione -446 (definendo posizione +1 il primo nucleotide del codone ATG di inizio della traduzione), il motivo CARGCW8GAT, una variante del motivo CARG, sito di legame tipico delle proteine MADS-box (de Folter e Angenent, 2006). La sequenza CARGCW8GAT funge da sito di legame per la proteina MADS-box AGL15 (*AGAMOUS-like 15*) (de Folter e Angenent, 2006). Il motivo CARGCW8GAT è presente in tutte le sequenze analizzate, ma non nel promotore del gene *PI* di *A. thaliana*. E' ipotizzabile quindi che tale motivo possa svolgere un ruolo regolativo critico nelle regioni promotrici degli orologi del locus *GLO-like* delle monocotiledoni.

In tutte le sequenze analizzate sono stati individuati siti di legame per i fattori di trascrizione AG e AG-like 3. E' noto che in *A. thaliana* l'eterotetramero AG/AP3/PI/SEP3 promuove e stabilizza, tramite un meccanismo autoregolativo, l'espressione tardiva di *PI* (Honma e Goto, 2001; Jack, 2001; Gomez-Mena et al., 2005). E' ipotizzabile che i siti di legame per AG identificati nel

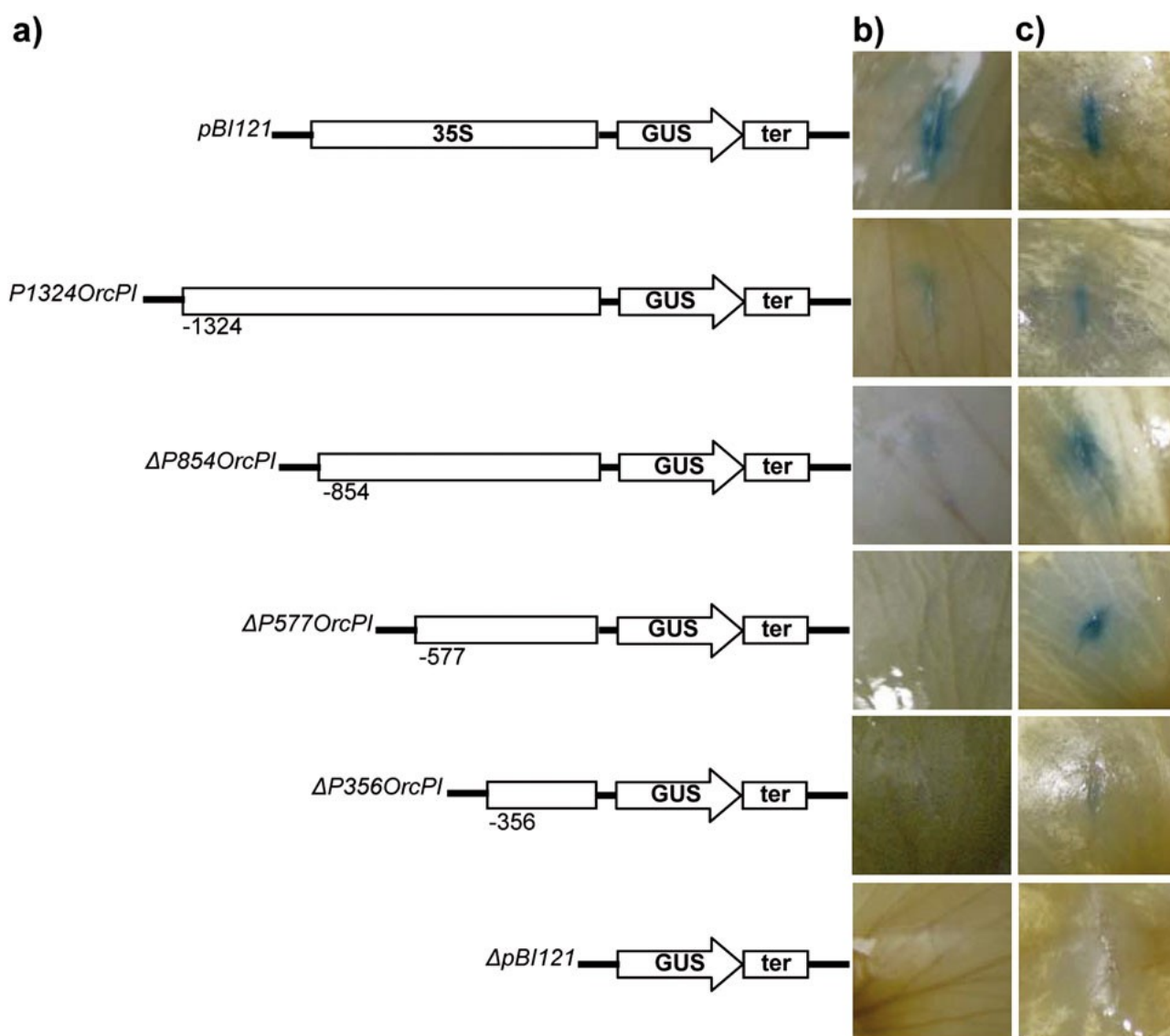


Figura 33 - Risultati del saggio GUS. a) Schema dei costrutti utilizzati, inclusi il controllo negativo ($\Delta pBI121$) e il controllo positivo ($pBI121$). b) Risultati degli esperimenti di trasformazione transiente in petali di rosa bianca dopo due e c) cinque giorni di incubazione.

presente lavoro rappresentino il bersaglio del complesso multiproteico AG/PI/AP3/SEP3 durante l'espressione auto-regolativa dei geni *GLO-like*.

Tramite l'analisi condotta con il software FootPrinter sono stati individuati nei promotori di *OrcPI*, *ORY1*, *ORY5* e *LIL* tre motivi presenti anche nel promotore prossimale di del gene *PI* di *A. thaliana*. Uno di questi motivi corrisponde all'elemento negativo (NE) (Chen et al., 2000) del promotore del gene *PI* di *A. thaliana*; inoltre, nel promotore del gene *OrcPI* è stato identificato un motivo ripetuto (in posizione -774), presente anche nelle sequenze omologhe di *L. regale* (in posizione -559) e di *O. sativa* (in posizione -7 a -32). E' possibile che tali sequenze ripetute (simili a microsatelliti imperfetti) svolgano un ruolo regolativo dell'espressione genica fungendo da siti di legame per i fattori di trascrizione e/o influenzando la struttura secondaria del DNA (Fujimori et al., 2003; Iglesias et al., 2004) (Figura 34, Tabella 14).





















Caratterizzazione strutturale e variabilità nucleotidica del gene *OrcPI*

Tramite PCR è stata amplificata, clonata e sequenziata la regione genomica del locus *OrcPI* che si estende dal promotore al sesto esone e la regione compresa tra il settimo esone e la 3'-UTR in *O. italica* ed in altre 12 specie di orchidee (Tabella 1). Nonostante i numerosi tentativi, non è stato possibile amplificare il sesto introne (potrebbe essere un introne molto lungo e/o contenere sequenze nucleotidiche che favoriscono la formazione di strutture secondarie che inibiscono la reazione di amplificazione) (Figura 35).

Il locus *OrcPI* comprende 7 esoni e 6 introni, organizzazione genomica tipica dei geni MADS-box di classe B, con dimensioni degli esoni e posizioni degli introni molto conservate.

Le sequenze nucleotidiche di *OrcPI* nelle 13 specie analizzate sono state allineate per determinare il livello di variabilità interspecifica delle regioni codificanti e non codificanti. Gli introni presentano una maggior variabilità nucleotidica (320 polimorfismi su 552 siti allineati, 57,8%) rispetto alle regioni

Legenda della figura 34

ORCPI -48 tccctactg	ORCPI -54 tccccctccc	ORCPI -77 ttttcttcga	ORCPI -112 cttcttcctc
 LILGLOB -396 tccctcctg	 LILGLOB -891 tccccctcca	 LILGLOB -375 tttcttaga	LILGLOB -454 cctcttcac
ORY1 -304 tccctcctc	ORY1 -135 tctctctcca	ORY5 -73 tttcttcga	 ORY1 -164 cctcttcctc
ORY5 -197 tccctcctc	ORY5 -190 tctctctcca		ORY5 -195 cctcctcctc
			ORY5 -192 cctcctcctc
ORCPI -115 ctctcttcttc	ORCPI -138 ccttcctca	ORCPI -144 tcccaacctt	ORCPI -204 ctccccgtact
 LILGLOB -22 caccttcttc	LILGLOB -145 cctcctcctca	 LILGLOB -26 accccacctt	LILGLOB -393 ctctctggact
ORY1 -136 ctctcctcctcc	 ORY1 -303 cctcctcctca	ORY1 -300 tcctcacctt	 ORY1 -144 ctctctttact
ORY5 -194 ctctcctcctcc	ORY5 -229 cccccctca	ORY5 -119 tcccaacctt	ORY5 -173 ctctctttact
	ORY5 -196 cctcctcctcc		
ORCPI -440 atatgattg	ORCPI -557 atcatgccat	ORCPI -772 gagaaggag	ORCPI -809 gaagaagag
 LILGLOB -800 atattattg	LILGLOB -313 attatgctat	LILGLOB -559 gagaagaag	LILGLOB -557 gaagaagag
ORY5 -1039 atatgattg	ORY1 -451 atcatgctat	ORY1 -23 gaggaggag	 ORY1 -978 gaagaaggg
	ORY5 -781 attatgctat	ORY1 -20 gaggaggag	ORY5 -101 gaagaacgg
		ORY1 -17 gaggaggag	
		ORY1 -14 gaggaggag	
		ORY5 -21 gaggaggag	
		ORY5 -18 gaggaggag	
		ORY5 -15 gaggaggag	
ORCPI -833 tttgtgggtt	ORCPI -1017 ttccatggc	ORCPI -1018 cttccatgg	ORCPI -1020 tcctttccat
LILGLOB -383 tttgtgggtt	LILGLOB -503 ctccatggc	LILGLOB -827 ctgccatgg	LILGLOB -889 ccctttccat
 ORY1 -737 tttgtgggtt	 ORY1 -156 tccccatggc	 ORY1 -157 ctccccatgg	 ORY1 -159 tcctccccat
ORY5 -1184 tttgtgatt	ORY5 -210 ttccatggc	ORY5 -211 cttccatgg	ORY5 -213 tcctttccat
ORY5 -1075 tttgggggtt			
ORY5 -1003 tttgggggtt			
ORCPI -1053 atcctctttct	ORCPI -1063 ccccaccat	ORCPI -1064 cccccacca	ORCPI -1070 ccaaagccccc
LILGLOB -456 atcctctttcc	ORCPI -143 cccaacctt	LILGLOB -488 ccccaacca	LILGLOB -84 cgaaagacccc
 ORY1 -136 ctctcctcctcc	 LILGLOB -25 ccccaacctt	ORY1 -125 cccccaccg	 ORY1 -186 cgaaagcctcc
ORY5 -194 ctctcctcctcc	ORY1 -299 cctcacctt	ORY5 -128 cccccacca	ORY5 -235 cgaaagccccc
ORY5 -191 ctctcctcctcc	ORY5 -118 ccccaacctt		
ORCPI -1106 gagatgaag	ORCPI -1194 agatgaaga	ORCPI -1297 ttttctatt	
 LILGLOB -559 gagaagaag	 LILGLOB -558 agaagaaga	LILGLOB -235 ttttctaatt	
ORY5 -1195 gagatgaag	ORY5 -1194 agatgaaga	 LILGLOB -220 ttttctaatt	
		ORY1 -672 ttttctctt	
		ORY5 -1237 ttttctctt	

Nome nella banca dati	Motivo consensus	Motivo di Footprinter	Sequenza in <i>OrcPI</i>
AG	NTTWCCWAAWNNGGNAAN	ccaaagcccc	caatCCAAAgccccacc
Dof2	NNNWAAAGCNN	ccaaagcccc	atccAAAGCcc
Dof3	NMNNAAGNNN	ccaaagcccc	atccAAAGCcc
ATHB-9	NNNNGTAATGATTRCNYBS	ccccaccat	ccccaccATGATcctcttc
ATHB-9	NNNNGTAATGATTRCNYBS	atcatgccat	ctcaattATCATgccatca
ATHB-9	NNNNGTAATGATTRCNYBS	atatgattg	agagaatATGATtggatt
O2	TCATTCCACGTAGATR	tttcttcga	TCATTttcttcgatcg

Tabella 14 - Corrispondenza tra i motivi identificati da FootPrinter e quelli presenti nelle banche dati specifiche per elementi regolativi vegetali.

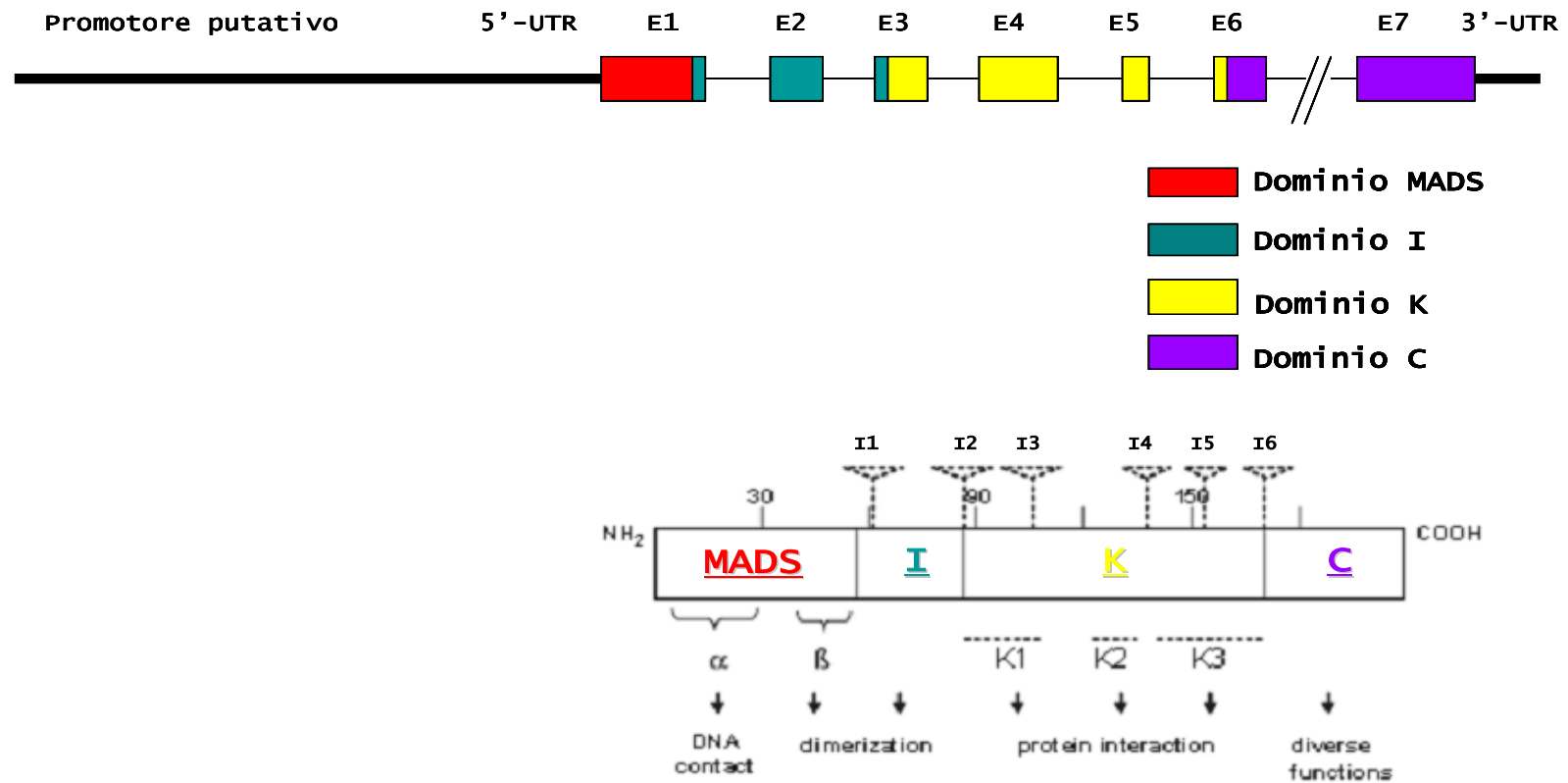


Figura 35 - Organizzazione genomica del locus *OrcPI* di *O. italica*.

regolative (promotore e UTRs) (119 siti variabili su 1516 allineati, 7,8%) e alla regione codificante del gene (9,6%). Il ridotto polimorfismo riscontrato nelle regioni regolative è in accordo con l'osservazione che i siti regolativi sono sottoposti a maggior costrizioni selettive rispetto alle regioni non regolative e quindi sono meno variabili, data la loro importanza nel regolare il processo di espressione genica.

Oltre all'analisi interspecifica è stata effettuata anche la stima del tasso di variabilità nucleotidica intraspecifica della regione che si estende dall'esone 1 all'esone 6 di *OrcPI* in 8 esemplari di *Anacamptis laxiflora* e 9 di *A. morio*. Il livello di divergenza nucleotidica tra le due specie è 0,095 ed è localizzato principalmente negli introni. Il polimorfismo riscontrato non è distribuito in modo uniforme lungo la regione analizzata: in *A. laxiflora* il polimorfismo dell'intera regione, stimato come valore di π , è di 0,00153 e la variabilità nucleotidica è ristretta alla regione che si estende dall'introne 3 all'introne 4; in *A. morio* il valore di π è 0,00718 e il polimorfismo si estende lungo tutta la regione. L'assenza di variabilità nucleotidica intraspecifica nell'introne 2 di entrambe le specie potrebbe indicare che questo introne abbia un tasso evolutivo inferiore rispetto a quello degli altri introni. Nonostante il test di Tajima condotto sulle regioni codificanti e non codificanti non indica significativa deviazione dalla neutralità, il test HKA è significativo ($P = 0,0415$) quando l'introne 1 è confrontato con la regione che comprende l'esone 2, l'introne 2 e l'esone 3. Questi risultati suggeriscono che le regioni non codificanti del gene *OrcPI* nelle orchidee sono soggette a costrizioni evolutive non omogenee, evidenziando un profilo selettivo eterogeneo dell'intero locus *OrcPI*, precedentemente descritto solo per i domini funzionali della regione codificante (Figura 36, Tabella 15) (Aceto et al. 2007).

Analisi dell'espressione del gene *OrcPI* in *O. italica*

L'analisi comparativa del profilo di espressione dei geni omologhi è uno strumento utile al chiarimento della funzione di tali geni e della loro evoluzione. L'espressione del gene *OrcPI* è stata esaminata in differenti tessuti di *O. italica*: radici,

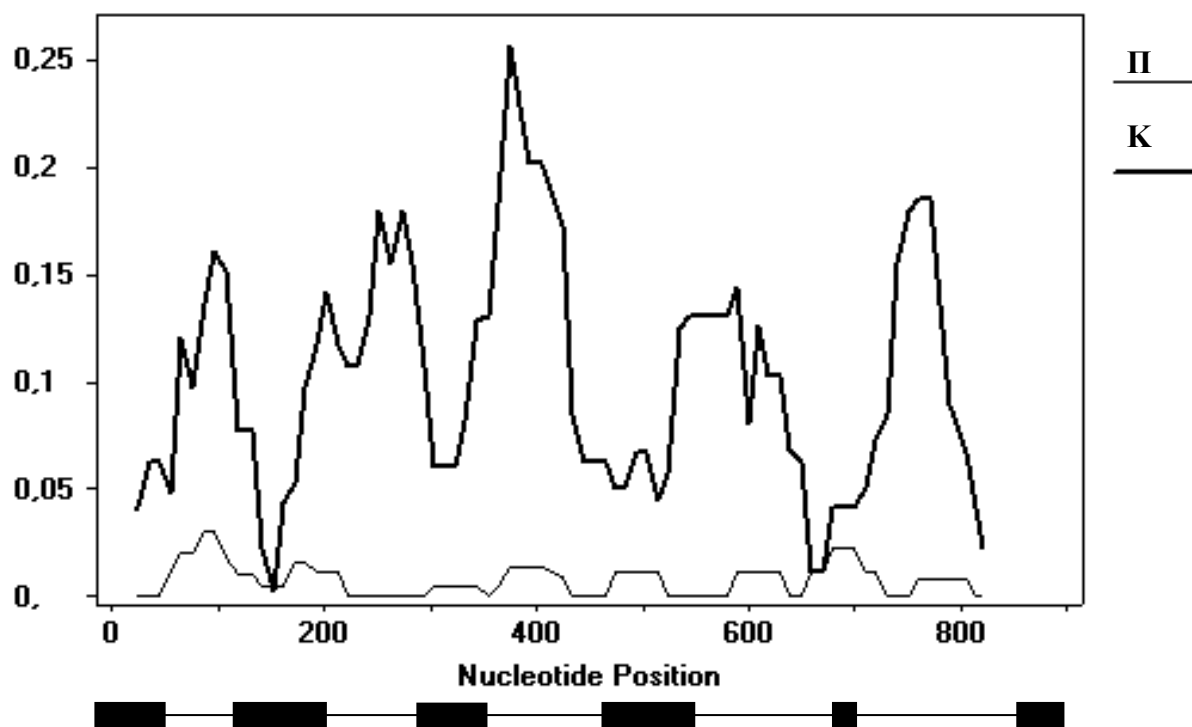


Figura 36 - Rappresentazione grafica della variabilità nucleotidica della regione di *OrcPI* che si estende dall'esone 1 all'esone 6 in *A. morio* e *A. laxiflora*. K indica la divergenza tra *A. laxiflora* e *A. morio*; π rappresenta il polimorfismo all'interno di *A. morio*.

Specie	Domini	Siti	S	n	Tajima's D
<i>A. laxiflora</i> (n = 8)	Exon 1	52	0	0.0000	n.d.
	Intron 1	78	0	0.0000	n.d.
	Exon 2	66	0	0.0000	n.d.
	Intron 2	83	0	0.0000	n.d.
	Exon 3	61	0	0.0000	n.d.
	Intron 3	91	1	0.0047	0.3335
	Exon 4	99	1	0.0043	0.3335
	Intron 4	102	1	0.0042	0.3335
	Exon 5	29	0	0.0000	n.d.
	Intron 5	133	0	0.0000	n.d.
	Exon 6	38	0	0.0000	n.d.
	Total	843	3	0.0015	0.4577
<i>A. morio</i> (n = 9)	Exon 1	52	0	0.0000	n.d.
	Intron 1	71	4	0.0211	0.0780
	Exon 2	66	2	0.0118	0.1959
	Intron 2	83	0	0.0000	n.d.
	Exon 3	61	1	0.0036	-1.0882
	Intron 3	80	3	0.0083	-1.5130
	Exon 4	99	1	0.0056	1.4012
	Intron 4	87	1	0.0064	1.4012
	Exon 5	29	0	0.0000	n.d.
	Intron 5	128	3	0.0117	1.3712
	Exon 6	38	0	0.0000	n.d.
	Total	805	15	0.0072	0.2262

Tabella 15 - Polomorfismo nucleotidico (Π) del locus *OrcPI* di *A. laxiflora* e *A. morio* (S, siti segreganti; n.d., non determinato).

foglie, infiorescenze in tre differenti stadi di sviluppo, tepali esterni, tepali interni, labello, ovario, colonna, fiori maturi e senescenti. I risultati mostrano che *OrcPI* non è espresso nei tessuti vegetativi ed è espresso in tutti gli organi fiorali, anche se a livelli più bassi nell'ovario; i livelli di mRNA presenti nel fiore maturo tendono a diminuire nel fiore in senescenza (Figura 37).

La presenza in *O. italica* di trascritti del gene *OrcPI* in tutti gli organi fiorali differisce dalla situazione riscontrata nelle dicotiledoni, nelle quali l'espressione dei geni di classe B è generalmente ristretta a petali e stami. Questo profilo di espressione potrebbe spiegare la similitudine morfologica di sepal e petali dei fiori di *O. italica*.

Il profilo di espressione di *OrcPI* è sovrapponibile a quello di *PeMADS6* di *Phalaenopsis equestris*, *PhPI15* di *P. hybrid cultivar*, *DcOPI* di *Dendrobium crumenatum* (Tsai et al., 2005, Xu et al., 2006, Guo et al., 2007) e a quello della copia *HrGLO1* di *H. radiata* (Kim et al., 2007).

Nelle orchidee *Dendrobium* e *Phalaenopsis*, i trascritti *PI/GLO-like* sono presenti nell'ovario immaturo e la loro concentrazione diminuisce dopo l'impollinazione, evento fondamentale per dare inizio alla maturazione dell'ovario nelle orchidee (Tsai et al. 2005, Xu et al., 2006). L'espressione di *OrcPI* nell'ovario di *O. italica*, sebbene a livelli inferiori rispetto alle altre strutture fiorali, ricalca la situazione generale riscontrata nelle altre orchidee. Sono necessari ulteriori studi, ad esempio conducendo esperimenti di impollinazione controllata, per stabilire se *OrcPI* sia coinvolto in modo diretto o indiretto nella repressione della maturazione dell'ovario. La persistenza dei trascritti di *OrcPI* nell'infiorescenza dall'antesi alla senescenza conferma il coinvolgimento dei geni *GLO-like* nella longevità fiorale delle orchidee, come descritto anche in *Phalaenopsis* (Tsai et al., 2005).

Nel laboratorio in cui ho svolto il lavoro di tesi è in corso l'analisi di espressione del gene *OrcPI2* in *O. italica* per confrontarlo con quello di *OrcPI* e chiarire il ruolo svolto dai due paraloghi nell'ontogenesi fiorale.

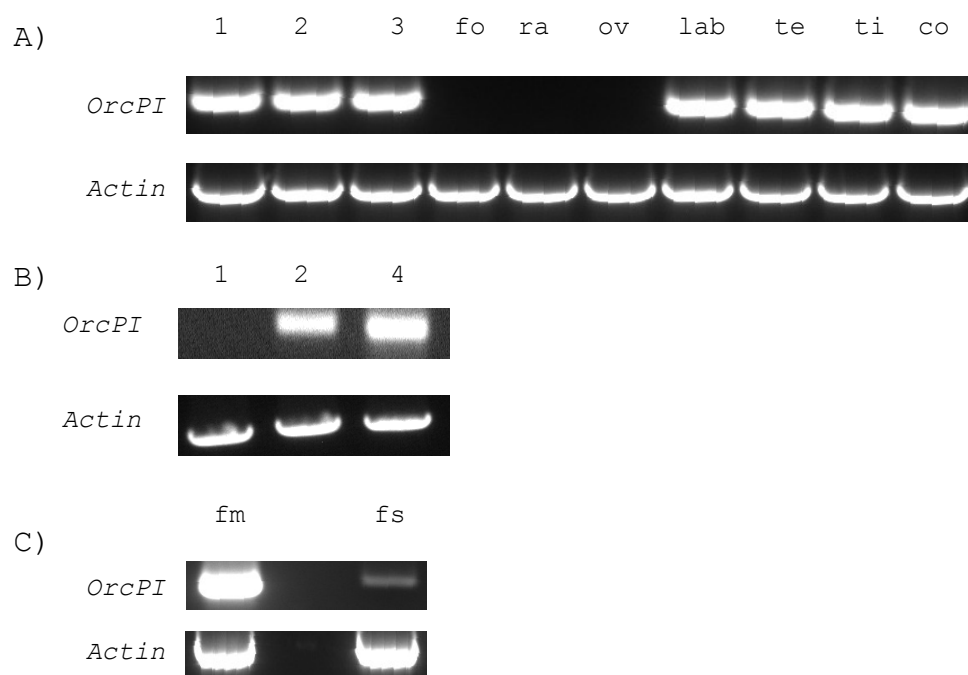


Figura 37 - Profilo di espressione del locus *OrcPI* in *O. italica* in (A) infiorescenze a tre differenti stadi di sviluppo (1 cm lunghezza x 0,5 cm altezza, 1,5 cm x 0,8 cm, 2,5 cm x 1,4 cm, definiti come stadio 1, 2, e 3, rispettivamente), foglie (fo), radici (ra), ovario (ov), labello (lab), tepali esterni (te) e interni (ti), colonna (co); (B) ovario con quantità crescenti di cDNA usati nella PCR (1 µl, 2 µl, e 4 µl della reazione di RT-PCR); (C) fiore maturo (fm, 6 giorni dopo la fioritura) e fiore senescente (fs, dopo 12 giorni dalla fioritura).

Bibliografia

- Aceto S., Caputo P., Cozzolino S., Gaudio L., Moretti A., 1999. **Phylogeny and evolution of *Orchis* and allied genera based on ITS DNA variation: morphological gaps and molecular continuity.** *Mol. Phylogenet. Evol.* 13: 67-76.
- Baum D., 1998. **The evolution of plant development.** *Curr. Opin. Plant Biol.* 1: 79-86.
- Becker A., Theißen G., 2003. **The major clades of MADS-box genes and their role in the development and evolution of flowering plants.** *Mol. Phylogenet. Evol.* 29: 464-489.
- Blanchette M., Tompa M., (2003). **FootPrinter: A program designed for phylogenetic footprinting.** *Nucleic Acids Res* 31: 3840-3842.
- Boffelli D., McAuliffe J., Ovcharenko D., Lewis K.D., Ovcharenko I., Pachter L., Rubin E.M., (2003). **Phylogenetic shadowing of primate sequences to find functional regions of the human genome.** *Science* 299:1391-1394.
- Buzgo M., Soltis D.E., Soltis P.S., 2004. **Floral developmental morphology of *Alborella trichopoda* (Amborellaceae).** *Int J. Plant Sci.* 165: 925-947.
- Carrol S., 2000. **Endless Forms Most Beautiful- the New Science of Evo Devo.** Norton & Co.
- Chanderbali A., 2006. **Genetic footprints of stamen ancestors guide perianth evolution in *Persea* (Lauraceae).** *Int. J. Plant Sci.* 167: 1075-1089.
- Chang W.C., Lee T.Y., Huang H.D., Huang H.Y., Pan R.L., (2008). **PlantPAN: Plant Promoter Analysis Navigator, for identifying combinatorial cis-regulatory elements with distance constraint in plant gene groups.** *BMC Genomics* 9: 561.
- Chen X., Riechmann, J.L., Jia D., Meyerowitz E., 2000. **Minimal regions in the *Arabidopsis PISTILLATA* promoter responsive to the APETALA3/PISTILLATA feedback control do not contain a CarG-box.** *Sex Plant Reprod.* 13: 85-94.
- Chung Y., Kim S., Kang H., Noh Y., Park M., Finkel D., An G., 1995. **Characterization of two rice MADS box genes homologous to GLOBOSA.** *Plant Sci.* 109: 45-56.
- Coen E.S., Meyerowitz E.M., 1991. **The war of whorls: genetic interactions controlling flower development.** *Nature* 353: 31-37.

- Cozzolino S., Aceto S., Caputo P., Gaudio L., Nazzaro R., 1998. **Phylogenetic relationship in *Orchis* and some related genera: an approach using chloroplast DNA.** *Nord. J. Bot.*, 18: 79-87.
- Cozzolino S., Widmer A., 2005. **Orchid diversity: an evolutionary consequence of deception?** *Trends in Ecology and Evolution* 20: 487-492.
- Davies B., DiRosa A., Eneva T., Saedler H., Sommer H., 1996. **Alteration of tobacco floral organ identity by expression of combinations of *Antirrhinum* MADS-box genes.** *Plant J.* 10: 663-677.
- De Bodt S., Theißen G., Van de Peer Y., 2006. **Promoter analysis of MADS-box genes in eudicots through phylogenetic footprinting.** *Mol. Biol. Evol.* 23(6): 1293-1303.
- De Folter S., Angenent G.C., 2006. **Trans meets cis in MADS science.** *Trends in Plant Science* 11: 224-231.
- Fang F., Blanchette M., (2006). **FootPrinter3: phylogenetic footprinting in partially alignable sequences.** *Nucleic Acids Res* 34: 617-620.
- Freeling M., Thomas B.C., 2006. **Gene-balanced duplications, like tetraploidy, provide predictable drive to increase morphological complexity.** *Genome Res.* 16: 805-814.
- Friis M., 2006. **Cretaceous angiosperm flower: innovation and evolution in plant reproduction.** *Paleogeogr., Paleoclimatol. Paleoecol.* 232: 251-293.
- Fujimori S., Washio T., Higo K., Ohtomo Y., Murakami K., Matsubara K., Kawai J., Carninci P., Hayashizaki Y., Kikuchi S., Tomita M., (2003). **A novel feature of microsatellites in plants: a distribution gradient along the direction of transcription.** *FEBS Lett* 554: 17-22.
- Gomez-Mena C., de Folter S., Costa M.M., Angenent G.C., Sablowski R., (2005). **Transcriptional program controlled by the floral homeotic gene *AGAMOUS* during early organogenesis.** *Development* 132: 429-438.
- Gumucio D.L., Heilstedt-Williamson H., Gray T.A., Tarle S.A., Shelton D.A., Tagle D.A., Slightom J.L., Goodman M., Collins F.S., (1992). **Phylogenetic footprinting reveals a nuclear protein which binds to silencer sequences in the human gamma and epsilon globin genes.** *Mol. Cell Biol.* 12: 4919-4929.

- Guo B., Hexige S., Zhang T., Pittman J.K., Chen D., Ming F., (2007). **Cloning and characterization of a PI-like MADS-Box gene in *Phalaenopsis* orchid.** *J Biochem Mol Biol* 40: 845-852.
- Hall T.A., (1999). **BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.** *Nucl Acids Symp Ser* 41: 95-98.
- Hernández-Hernández T., Martínez-Castilla L.P., Alvarez-Buylla E.R., (2007). **Functional diversification of B MADS-box homeotic regulators of flower development: adaptive evolution in protein-protein interaction domains after major gene duplication events.** *Mol. Biol. Evol.* 24: 465-481.
- Higo K., Ugawa Y., Iwamoto M., Korenaga T., (1999). **Plant cis-acting regulatory DNA elements (PLACE) database: 1999.** *Nucleic Acids Res* 27: 297-300.
- Hong R.L., Hamaguchi L., Busch M.A., Weigel D., (2003). **Regulatory elements of the floral homeotic gene AGAMOUS identified by phylogenetic footprinting and shadowing.** *Plant Cell* 15: 1296-1309.
- Honma T., Goto K., (2000). **The Arabidopsis floral homeotic gene PISTILLATA is regulated by discrete cis-elements responsive to induction and maintenance signals.** *Development* 127: 2021-2030.
- Honma T., Goto K., (2001). **Complexes of MADS-box proteins are sufficient to convert leaves into floral organs.** *Nature* 409: 525-529.
- Hsu H., Yang C., 2002. **An orchid (*Oncidium Gower Ramsey*) AP3-like MADS genes regulate floral formation and initiation.** *Plant Cell Physiol.* 43: 1198-1209.
- Hudson R.R., Kreitman M., Aguade M., 1987. **A test of neutral molecular evolution based on nucleotide data.** *Genetics* 116: 153-9.
- Iglesias A.R., Kindlund E., Tammi M., Wadelius C., (2004). **Some microsatellites may act as novel polymorphic cis-regulatory elements through transcription factor binding.** *Gene* 341: 149-165.
- Inada D.C., Bashir A., Lee C., Thomas B.C., Ko C., Goff S.A., Freeling M., (2003). **Conserved noncoding sequences in the grasses.** *Genome Res* 13: 2030-2041.
- Jack T., (2001). **Relearning our ABCs: new twists on an old model.** *Trends Plant Sci* 6: 310-316.

Jefferson R., Kavanagh T., Bevan M., 1987. **GUS fusion: β -glucuronidase as a sensitive and versatile gene fusion marker in higher plants.** *The EMBO Journal* 6: 3901-3907.

Jersakova J., Johnson S., Kindlmann P., 2006. **Mechanisms and evolution of deceptive pollination in orchids.** *Biological Reviews Of the Cambridge Philosophical Society* 81: 219-235.

Johnson S., Linder H., Steiner K., 1998. **Phylogeny and radiation of pollination system in *Disa* (Orchidaceae).** *American Journal of Botany* 85: 402-411.

Jongmin N., Joonyul K., Shinyoung L., Gynheung A., Hong M., Masatoshi N., 2003. **Type I MADS-box genes have experienced faster birth-and-death evolution than type II MADS-box genes in angiosperms.** *PNAS* 101: 1910-1915.

Jukes T.H. e Cantor C.R., 1969. **Evolution of protein molecules.** In Munro HN, editor, *Mammalian Protein Metabolism*, pp. 21-132, Academic Press, New York.

Kaufmann K., Melzer R., Theißen G., 2005. **MIKC-type MADS-domain proteins: structural modularity, protein interactions and network evolution in land plants.** *Gene* 347: 183-198.

Kim S., Yoo M.J., Albert V.A., Farris J.S., Soltis P.S., Soltis D.E., 2004. **Phylogeny and diversification of B-function MADS-box genes in angiosperms: evolutionary and functional implications of a 260-million-year-old duplication.** *Am. J. Bot.* 91: 2102-2118.

Kim S., Koh J., Ma H., Hu Y., Endress P.K., Buzgo M., Hauser B.A., Soltis P.S., Soltis D.E., 2005. **Sequence and expression studies of A-, B-, and E-class MADS-box genes in *Eupomatia* (Eupomatiaceae): support for the bractate origin of the caliptra.** *Int. J. Plant Sci.* 166: 185-198.

Kim S.Y., Yun P.Y., Fukuda T., Ochiai T., Yokoyama J., Kameya T., Kanno A., (2007). **Expression of a *DEFICIENS*-like gene correlates with the differentiation between sepal and petal in the orchid, *Habenaria radiata* (Orchidaceae).** *Plant Science* 172: 319-326.

Kramer E.M., Dorit R.L., Irish V.F., 1998. **Molecular evolution of genes controlling petal and stamen development: duplication and divergence within the *APETALA3* and *PISTILLATA* MADS-box gene lineages.** *Genetics* 149: 765-783.

Kramer E.M., Jaramillo M., Di Stilio V.S., 2004. **Patterns of gene duplication and functional evolution during the diversification of the AGAMOUS subfamily of MADS box genes in angiosperms.** *Genetics* 166: 1011-1023.

Kumar S., Tamura, K., Nei M., 2004. **Prospects for inferring very large phylogenies by using the neighbor-joining method.** *Proc. Natl. Acad. Sci. U.S. A.* 101: 11030-11035.

Lawton-Rauh A.L., Alvarez-Buylla E.R., Purugganam M.D., 2000. **Molecular evolution of flower development.** *Trends ecol. Evol.* 15: 144-149.

Lescot M., Dehais P., Thijs G., Marchal K., Moreau Y., Van de Peer Y., Rouze P., Rombauts S., (2002). **PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences.** *Nucleic Acids Res* 30: 325-327.

Li Z.T., Gray D.J., 2005. **Isolation by improved thermal asymmetric interlaced PCR and characterization of a seed-specific 2S albumin gene and its promoter from grape (*Vitis vinifera* L.).** *Genome* 48: 312-320.

Litt A., Irish V.F., 2003. **Duplication and diversification in the APETALA1/FRUITFULL floral homeotic gene lineage: implications for the evolution of floral development.** *Genetics* 165: 821-833.

Liu Y., Whittier R.F., 1995. **Thermal Asymmetric Interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking.** *Genomics* 25: 674-681.

Malcomber S.T., Kellog E.A., 2005. **SEPALLATA gene diversification: Brave new whorls.** *Trends in Plant Science* 10: 427-435.

Martinez-Castilla L.P., Alveraz-Buylla E.R., 2003. **Adaptative evolution in the Arabidopsis MADS-box gene family inferred from its complete resolved phylogeny.** *Proc Natl Acad USA* 100: 13407-13412.

Messenguy F., Dubois E., 2003. **Role of MADS-box proteins and their cofactors in combinatorial control of gene expression and cell development.** *Gene* 316: 1-21.

Mondragon-Palomino M., Hiese L., Harter A., Koch M.A., Theißen G., (2009). **Positive selection and ancient duplications in the**

evolution of class B floral homeotic genes of orchids and grasses. *BMC Evol Biol* 9: 81.

Mondragon-Palomino M., Theißen G., 2008. **MADS about the evolution of orchid flowers.** *Trends Plant Sci* 13: 51-59.

Mondragon-Palomino M., Theißen G., 2009. **Why are orchid so diverse? Reduction of evolutionary constraints by paralogues of class B floral homeotic genes.** *Annals of Botany* 104 (3): 583-594.

Moon Y., Jung J., Kang H., An G., 1999. **Identification of a rice APETALA 3 homologue by yeast two-hybrid screening.** *Plant Mol. Biol.* 40: 7076-7081.

Nadeau J., Zhang X., Li J., Oneill S., 1996. **Ovule development: identification of stage-specific and tissue-specific cDNAs.** *Plant Cell* 8: 213-239.

Nei M., Gojobori T., (1986). **Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions.** *Mol Biol Evol* 3: 418-26.

Nei M., 2005. **Selectionism and neutralism in molecular evolution.** *Mol. Biol. Evol.* 22: 2318-2342.

Nei M., Rooney A.P., 2005. **Concerted and birth-and-death evolution of multigene families.** *Annu. Rev. Genet.* 39: 121-152.

Notredame C., Higgins D.G., Heringa J. 2000. **T-Coffee: A Novel Method for Fast and Accurate Multiple Sequence Alignment.** *J. Mol. Biol.* 302: 205-217.

Okada K., Shimura Y., 1994. **Genetic analysis of signaling in flower development using Arabidopsis.** *Plant Mol. Biol.*, 26: 1357-77.

Pelaz S., Gustafson-Brown C., Kohalmi S.E., Crosby W.L., Yanofsky M.F., 2001. **APETALA1 and SEPALLATA3 interact to promote flower development.** *Plant J.* 26: 385-394.

Pinyopich A., Ditta D.S., Savidge B., Liljetgren S.J., Baumann E., Wisman E., Yanofsky M.F., 2003. **Assessing the redundancy of MADS-box genes during carpel and ovule development.** *Nature* 424: 85-88.

Purugganam M.D., Suddith J.I., 1998. **Molecular population genetics of thre Arabidopsis CAULIFLOWER regulatory gene: non-neutral evolution and wild variation in floral homeotic function.** *Proceedings of the National Accademy of Sciences of the USA* 95: 8130-8134.

Ramirez S., Gravendeel B., Singer R., Marshall C., Pierce N., 2007. **Dating the origin of the Orchidaceae from a fossil orchid with its pollinator.** *Nature* 448: 1042-1045.

Riechmann J.L., Krizek B.A., Meyerowitz E.M., 1996. **Dimerization specificity of *Arabidopsis* MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA, and AGAMOUS.** *Proc. Natl. Acad. Sci. U.S.A.* 93: 4793-4798.

Rijpkema A.S., Gerats T., Vandenbussche M., 2007. **Evolutionary complexity of MADS complexes.** *Curr. Opin. in Plant Bio.* 10: 32-38.

Robinson-Rechavi M., Huchon D., 2000. **RRTree: relative-rate tests between groups of sequences on a phylogenetic tree.** *Bioinformatics* 16: 296-297

Rozas J., Sánchez-DelBarrio J.C., Messeguer X., Rozas R., (2003). **DnaSP, DNA polymorphism analyses by the coalescent and other methods.** *Bioinformatics* 19: 2496-2497.

Rudall P., Bateman R., 2002. **Roles of synorganisation, zygomorphy and heterotopy in floral evolution: the gynostemium and labellum of orchids and other liloid monocots.** *Biological reviews of the Cambridge Philosophical Society* 77: 403-441.

Shan H., Zahn L., Guindon S., Wall K.P., Kong H., Ma H., dePamphilis C.W., Leebens-Mack J., 2009. **Evolution of plant MADS Box transcription factors: evidence for shifts in selection associated with early angiosperm diversification and concerted gene duplications.** *Mol. Biol. and Evol.* 26 (10): 2229-2244.

Shan H., Zhang N., Liu C., Xu G., Zhang J., Chen Z., Kong H., 2007. **Patterns of gene duplication and functional diversification during the evolution of the AP1/SQUA subfamily of plant MADS-box genes.** *Mol. Phylogenet. Evol.* 44: 26-41.

Skipper M., Johansen L., Pedersen K., Frederiksen S., Johansen B., 2006. **Cloning and transcription analysis of an AGAMOUS- and SEEDSTICK orthologues in the orchid *Dendrobium thyrsiflorum*.** *Gene* 366: 266-274.

Soltis D.E., Ma H., Frohlich M.W., Soltis P.S., Albert V.A., Oppenheimer D.G., Altman N.S., dePamphilis C., Leebens-Mack J., 2007. **The floral genome: an evolutionary history of gene duplication and shifting patterns of gene expression.** *Trends Plant Sci.* 12: 358-367.

- Song I., Nakamura T., Fukuda T., 2006. **Spatiotemporal expression of duplicate AGAMOUS orthologues during floral development in Phalaenopsis.** *Dev. Genes and Evol.* 216: 301-313.
- Tajima F., 1989. **Statistical method for testing the neutral mutation hypothesis by DNA polymorphism.** *Genetics* 123: 585-95.
- Tamura K., Dudley J., Nei M., Kumar S., (2007). **MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0.** *Mol Biol Evol* 24: 1596-9.
- Terauchi R., Kahl G., 2000. **Rapid isolation of promoter sequences by TAIL-PCR: the 5'-flanking regions of *Pal* and *Pgi* genes from yams (*Dioscorea*).** *Mol. Gen. Genet.* 263: 554-560.
- Theißen G., 2001. **Development of floral organ identity: stories from the MADS house.** *Curr. Opin. Plant Biol.* 4: 75-85.
- Tsai W., Kuoh C., Chuang M., Chen W., Chen H., 2004. **Four DEF-like MADS-box genes displayed distinct floral morphogenetic roles in Phalaenopsis orchid.** *Plant Cell Physiol.* 45: 831-844.
- Tsai W., Lee P., Chen H., 2005. **PeMADS6, a GLOBOSA/PISTILLATA-like gene in Phalaenopsis equestris involved in petaloid formation and correlated with flower longevity and ovary development.** *Plant cell Physiology* 46: 1125-1139.
- Vandenbussche M., Theißen G., Van de Peer Y., Gerats T., 2003. **Structural diversification and neo-functionalization during floral MADS box gene evolution by C-terminal frameshift mutations.** *Nucleic Acids Res.* 31: 4401-4409.
- Winter K., Weiser C., Kaufmann K., Böhne A., Kirchner C., Kanno A., Saedler H., Theißen G., 2002. **Evolution of class B Floral homeotic proteins: obligate heterodimerization originated from homodimerization.** *Mol Biol. Evol.* 19 (5): 587-596.
- Xu G., Kong H., 2007. **Duplication and divergence of floral MADS-box genes in grasses: evidence for the generation and modification of novel regulators.** *J. Integr. Plant Biol.* 49: 927-939.
- Xu Y., Teo L., Zhou J., Kumar P., Yu H., 2006. **Floral organ identity genes in the orchid *Dendrobium crumenatum*.** *The Plant Journal* 46: 54-68.
- Yang Z., 1997. **PAML: a program package for phylogenetic analysis by maximum likelihood.** *CABIOS* 13: 555-556.

Yang Z., Nielsen R., Goldman N., Pedersen A.M.K., 2000. **Codon substitution models for heterogeneous selection pressure at amino acid sites.** *Genetic* 155: 431-441.

Yang Z., Nielsen R., Goldman N., Pedersen A.M.K., 2005. **PAML: a phylogenetic analysis by a maximum likelihood.** 3.15 *th edition*.

Yu H., Ito T., Wellmer F., Meyerowitz E., 2004. **Repression of AGAMOUS-like 24 is a crucial step in promoting flower development.** *Nat. Genet.* 36: 225-240.

Zahn L.M., Kong H., Leebens-Mack J.H., Kim S., Soltis P.S., Landherr L.L., Soltis D.E., Depamphilis C.W., Ma H., 2005. **The evolution of the SEPALLATA subfamily of MADS-box genes: a preangiosperm origin with multiple duplications throughout angiosperm history.** *Genetics* 169: 2209-2223.

Zhang X., O'Neill S., 1993. **Ovary and gametophyte development are coordinately regulated by auxin and ethylene following pollination.** *Plant cell* 5: 403-418.

Zhang S., Nielsen R., Yang Z., 2005. **Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level.** *Mol. Biol. Evol.* 12: 2472-2477.